GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: May 16, 2002, 14:58:45 ; Search time 6516.55 Seconds
(without alignments)
8901.698 Million cell updates/sec

Title: US-09-898-556A-3
Perfect score: 2772
Sequence: 1 caggcgcgttaagctggttg.....tttctacccatcctcaccct 2772

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1797656 seqs, 10463268293 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2772	100.0	2772	9	AB013897	AB013897 Homo sapi
2	2383	86.0	2854	9	AK056243	AK056243 Homo sapi
c 3	2244	81.0	135173	9	AC008806	AC008806 Homo sapi
4	2244	81.0	253217	2	AC016590	AC016590 Homo sapi
5	2193	79.1	38173	9	AC093235	AC093235 Homo sapi
6	1771	63.9	1998	9	BC004513	BC004513 Homo sapi
7	672	24.2	672	9	HUMHKR1	M20675 Human Krupp
8	571	20.6	678	6	AX067361	AX067361 Sequence
9	37	1.3	130351	2	AC024911	AC024911 Homo sapi
10	37	1.3	161625	2	AC096898	AC096898 Homo sapi
11	37	1.3	164264	2	AC034197	AC034197 Homo sapi
12	37	1.3	165379	2	AC069265	AC069265 Homo sapi
13	37	1.3	184375	2	AC026219	AC026219 Homo sapi
14	36	1.3	148270	2	AC027579	AC027579 Homo sapi
15	36	1.3	174405	2	AC022727	AC022727 Homo sapi
16	36	1.3	179792	2	AP001531	AP001531 Homo sapi
17	34	1.2	150794	9	AC025211	AC025211 Homo sapi
c 18	34	1.2	166650	2	AC087493	AC087493 Homo sapi
19	34	1.2	218922	2	AC008106	AC008106 Homo sapi
c 20	32	1.2	162453	2	AC027607	AC027607 Homo sapi
c 21	32	1.2	171279	2	AC036185	AC036185 Homo sapi
c 22	32	1.2	171601	2	AC026269	AC026269 Homo sapi
c 23	32	1.2	173133	2	AC016166	AC016166 Homo sapi
24	32	1.2	174701	2	AC098581	AC098581 Homo sapi
25	32	1.2	175833	2	AC073412	AC073412 Homo sapi
c 26	32	1.2	207362	9	AC007380	AC007380 Homo sapi
27	31	1.1	154746	9	AL157944	AL157944 Human DNA
c 28	31	1.1	168736	2	AC019042	AC019042 Homo sapi
c 29	31	1.1	177864	9	AC024996	AC024996 Homo sapi
30	29	1.0	2328	9	HSM800283	AL049942 Homo sapi
31	29	1.0	2790	9	AK056466	AK056466 Homo sapi
32	29	1.0	3178	9	AK056184	AK056184 Homo sapi
33	29	1.0	3273	9	BC021298	BC021298 Homo sapi
c 34	29	1.0	127867	9	HS694B14	AL031673 Human DNA
35	29	1.0	170129	2	AC013770	AC013770 Homo sapi
c 36	29	1.0	181770	2	AC109473	AC109473 Homo sapi
37	28	1.0	2823	9	BC011862	BC011862 Homo sapi
38	28	1.0	3055	9	AK055734	AK055734 Homo sapi
39	28	1.0	61452	2	AC091936	AC091936 Homo sapi
c 40	28	1.0	62002	2	AF161800	AF161800 Homo sapi
41	28	1.0	65236	2	AC105030	AC105030 Homo sapi
42	28	1.0	80846	9	AL512489	AL512489 Human DNA
43	28	1.0	90220	9	HSJ734P14	AL049650 Human DNA
c 44	28	1.0	109217	9	AF159227	AF159227 Homo sapi
45	28	1.0	131328	9	AC099798	AC099798 Homo sapi

ALIGNMENTS

RESULT 1
AB013897
LOCUS AB013897 2772 bp mRNA linear PRI 03-NOV-1999
DEFINITION Homo sapiens mRNA for HKR1, partial cds.
ACCESSION AB013897
VERSION AB013897.1 GI:6177784
KEYWORDS HKR1.
SOURCE Homo sapiens leukemia cell cell_line:CMK86 cDNA to mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Oguri,T., Katoh,O., Takahashi,T., Isobe,T., Kuramoto,K., Hirata,S., Yamakido,M. and Watanabe,H.
TITLE The Kruppel-type Zinc Finger Family Gene, HKR1, Is Induced in Lung Cancer by Exposure to Platinum Drugs
JOURNAL Unpublished (1998)
REFERENCE 2 (bases 1 to 2772)
AUTHORS Katoh,O.
TITLE Direct Submission
JOURNAL Submitted (13-MAY-1998) Osamu Katoh, Res. Inst. Radiation Biology and Medicine, Hiroshima Univ., Department of Environment and Mutation; Kasumi 1-2-3, Minami-Ku,, Hiroshima, Hiroshima 734-8553,

```
          Japan (E-mail:katoho@mcai.med.hiroshima-u.ac.jp,
          Tel:+81-82-257-5817, Fax:+81-82-256-7104)
FEATURES             Location/Qualifiers
     source          1. .2772
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="19"
                     /map="19q12-19q13.1"
                     /cell_line="CMK86"
                     /cell_type="leukemia cell"
     gene            1. .2096
                     /gene="HKR1"
     CDS             <1. .2096
                     /gene="HKR1"
                     /function="transcription factor"
                     /note="The HKR1 protein is similar to Kruppel-type zinc
                     finger transcription factor and has Kruppel-associated box
                     (KRAB);-A and -B box at the N-terminus(nucleotides
                     193-390) followed by 13 C2H2-type zinc finger domain at
                     the C-terminus(1020-2084)."
                     /codon_start=3
                     /product="HKR1"
                     /protein_id="BAA86058.1"
                     /db_xref="GI:6177785"
                     /translation="GALSWLGPGKASLLRSFPHLCSLLPDFRLQDPRRAPAFHLSSET
                     LPFSRKSTQETRKMATGLLRAKKEAFVAFRDVAVYFTQEEWRLLSPAQRTLHREVMLE
                     TYNHLVSLEIPSSKPKLIAQLERGEAPWREERKCPLDLCPESKPEIQLSPSCPLIFSS
                     QQALSQHVWLSHLSQLFSSLWAGNPLHLGKHYPEDQKQQQDPFCFSGKAEWIQEGEDS
                     RLLFGRVSKNGTSKALSSPPEEQQPAQSKEDNTVVDIGSSPERRADLEETDKVLHGLE
                     VSGFGEIKYEEFGPGFIKESNLLSLQKTQTGETPYMYTEWGDSFGSMSVLIKNPRTHS
                     GGKPYVCRECGRGFTWKSNLITHQRTHSGEKPYVCKDCGRGFTWKSNLFTHQRTHSGL
                     KPYVCKECGQSFSLKSNLITHQRAHTGEKPYVCRECGRGFRQHSHLVRHKRTHSGEKP
                     YICRECEQGFSQKSHLIRHLRTHTGEKPYVCTECGRHFSWKSNLKTHQRTHSGVKPYV
                     CLECGQCFSLKSNLNKHQRSHTGEKPFVCTECGRGFTRKSTLITHQRTHSGEKPFVCA
                     ECGRGFNDKSTLISHQRTHSGEKPFMCRECGRRFRQKPNLFRHKRAHSGAFVCRECGQ
                     GFCAKLTLIKHQRAHAGGKPHVCRECGQGFSRQSHLIRHQRTHSGEKPYICRKCGRGF
                     SRKSNLIRHQRTHSG"
     polyA_site      2772
                     /note="11 a nucleotides"
BASE COUNT      746 a    671 c    720 g    635 t
ORIGIN
```

Query Match 100.0%; Score 2772; DB 9; Length 2772;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2772; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 caggcgcgttaagctggttgggacccgggaaggcctccctcttaaggtctttcccacacc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CAGGCGCGTTAAGCTGGTTGGGACCCGGGAAGGCCTCCCTCTTAAGGTCTTTCCCACACC 60

Qy      61 tctgctccttgttacctgactttcggcttcaggatccgcggcgtgcacccgcgttccatc 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TCTGCTCCTTGTTACCTGACTTTCGGCTTCAGGATCCGCGGCGTGCACCCGCGTTCCATC 120

Qy     121 tgtcttctgagactttgcccttctccaggaagagcactcaggagaccaggaaaatggcta 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 TGTCTTCTGAGACTTTGCCCTTCTCCAGGAAGAGCACTCAGGAGACCAGGAAAATGGCTA 180

Qy     181 cagggctcctgagagccaaaaaagaggcgttcgtggcattcagggatgtggctgtgtact 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 CAGGGCTCCTGAGAGCCAAAAAAGAGGCGTTCGTGGCATTCAGGGATGTGGCTGTGTACT 240

Qy     241 tcacccaggaggagtggaggttgttgagccctgctcagaggaccctgcacagggaggtga 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 TCACCCAGGAGGAGTGGAGGTTGTTGAGCCCTGCTCAGAGGACCCTGCACAGGGAGGTGA 300

Qy     301 tgctggagacttataaccatctggtctcactggaaattccatcttctaaaccaaaaactca 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 TGCTGGAGACTTATAACCATCTGGTCTCACTGGAAATTCCATCTTCTAAACCAAAACTCA 360

Qy     361 ttgctcagctggagcgaggggaagcgccctggagagaggagagaaaatgtccactggacc 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 TTGCTCAGCTGGAGCGAGGGGAAGCGCCCTGGAGAGAGGAGAGAAAATGTCCACTGGACC 420

Qy     421 tctgtccagaatcgaagccagaaattcaacttagtccctcctgccctctgattttctcca 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 TCTGTCCAGAATCGAAGCCAGAAATTCAACTTAGTCCCTCCTGCCCTCTGATTTTCTCCA 480

Qy     481 gtcagcaagctctcagccaacatgtgtggctgagtcatctctctcagctgttttcaagtt 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 GTCAGCAAGCTCTCAGCCAACATGTGTGGCTGAGTCATCTCTCTCAGCTGTTTTCAAGTT 540

Qy     541 tatgggcaggaaatcctctccacctgggaaaacactatccagaagatcagaaacaacagc 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 TATGGGCAGGAAATCCTCTCCACCTGGGAAAACACTATCCAGAAGATCAGAAACAACAGC 600

Qy     601 aggatccattctgctttagtggcaaagcagaatggattcaagagggagaagactccagac 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 AGGATCCATTCTGCTTTAGTGGCAAAGCAGAATGGATTCAAGAGGGAGAAGACTCCAGAC 660

Qy     661 tcctgtttgggagagtaagcaaaaatggcacttcaaaggcactttccagcccacctgaag 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 TCCTGTTTGGGAGAGTAAGCAAAAATGGCACTTCAAAGGCACTTTCCAGCCCACCTGAAG 720

Qy     721 aacaacagccagcacagtccaaggaagacaacacagtggtggatatagggtccagccctg 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 AACAACAGCCAGCACAGTCCAAGGAAGACAACACAGTGGTGGATATAGGGTCCAGCCCTG 780

Qy     781 aacggagggcagatctagaggaaacagacaaagtattgcatggtttagaagtctcaggat 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 AACGGAGGGCAGATCTAGAGGAAACAGACAAAGTATTGCATGGTTTAGAAGTCTCAGGAT 840

Qy     841 ttggagaaatcaaatatgaagagtttgggccaggctttatcaaggagtcaaacctcctta 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 TTGGAGAAATCAAATATGAAGAGTTTGGGCCAGGCTTTATCAAGGAGTCAAACCTCCTTA 900

Qy     901 gcctccagaagacacaaactggggagacaccttacatgtacactgagtggggagacagct 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 GCCTCCAGAAGACACAAACTGGGGAGACACCTTACATGTACACTGAGTGGGGAGACAGCT 960

Qy     961 ttggcagtatgtcagtcctcatcaaaaacccaaggacacactctgggggaaagccttatg 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 TTGGCAGTATGTCAGTCCTCATCAAAAACCCAAGGACACACTCTGGGGGAAAGCCTTATG 1020

Qy    1021 tgtgcagggaatgtgggcgaggctttacgtggaagtcaaacctgatcacacatcagagga 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 TGTGCAGGGAATGTGGGCGAGGCTTTACGTGGAAGTCAAACCTGATCACACATCAGAGGA 1080

Qy    1081 cacactcaggggagaaaccttatgtgtgcaaggattgtggacgaggctttacttggaagt 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 CACACTCAGGGGAGAAACCTTATGTGTGCAAGGATTGTGGACGAGGCTTTACTTGGAAGT 1140

Qy    1141 cgaacctctttacacatcagcggacacactcagggctcaagccttatgtgtgcaaggaat 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 CGAACCTCTTTACACATCAGCGGACACACTCAGGGCTCAAGCCTTATGTGTGCAAGGAAT 1200

Qy    1201 gtgggcagagctttagcctgaagtcaaacctcattacccaccagagggcgcacactgggg 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 GTGGGCAGAGCTTTAGCCTGAAGTCAAACCTCATTACCCACCAGAGGGCGCACACTGGGG 1260

Qy    1261 agaagccttatgtttgcagggaatgtgggcgtggctttcgccagcattcacacctggtca 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 AGAAGCCTTATGTTTGCAGGGAATGTGGGCGTGGCTTTCGCCAGCATTCACACCTGGTCA 1320

Qy    1321 gacacaagaggacacattcaggagagaagccttacatttgcagggagtgtgagcaaggct 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 GACACAAGAGGACACATTCAGGAGAGAAGCCTTACATTTGCAGGGAGTGTGAGCAAGGCT 1380

Qy    1381 ttagccagaagtcacacctcatcagacacttaaggacacacacaggagagaagccttatg 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 TTAGCCAGAAGTCACACCTCATCAGACACTTAAGGACACACACAGGAGAGAAGCCTTATG 1440

Qy    1441 tatgcacagaatgtgggcgtcactttagctggaaatcaaacctcaaaacacaccagagga 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1441 TATGCACAGAATGTGGGCGTCACTTTAGCTGGAAATCAAACCTCAAAACACACCAGAGGA 1500
```

```
Qy   1501 cacactcaggggttaaaccttatgtctgcctggagtgcgggcagtgctttagcctgaagt 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 CACACTCAGGGGTTAAACCTTATGTCTGCCTGGAGTGCGGGCAGTGCTTTAGCCTGAAGT 1560

Qy   1561 caaaccttaacaaacaccagaggtcacacacgggggagaagccatttgtatgtacggagt 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 CAAACCTTAACAAACACCAGAGGTCACACACGGGGGAGAAGCCATTTGTATGTACGGAGT 1620

Qy   1621 gtgggcgaggctttacccggaaatcaaccctgatcacgcaccagaggacacactcagggg 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 GTGGGCGAGGCTTTACCCGGAAATCAACCCTGATCACGCACCAGAGGACACACTCAGGGG 1680

Qy   1681 agaagccatttgtatgtgctgagtgtggacgaggctttaatgataagtccaccctcattt 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 AGAAGCCATTTGTATGTGCTGAGTGTGGACGAGGCTTTAATGATAAGTCCACCCTCATTT 1740

Qy   1741 cacaccagaggacacattcaggggaaaagccttttatgtgcagggagtgtggcagaaggt 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 CACACCAGAGGACACATTCAGGGGAAAAGCCTTTTATGTGCAGGGAGTGTGGCAGAAGGT 1800

Qy   1801 ttcggcagaagcctaacctgtttaggcacaagagggcacactcaggtgcctttgtgtgca 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 TTCGGCAGAAGCCTAACCTGTTTAGGCACAAGAGGGCACACTCAGGTGCCTTTGTGTGCA 1860

Qy   1861 gggagtgtgggcaaggcttttgtgctaagttaactctcattaaacaccagagagcacacg 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 GGGAGTGTGGGCAAGGCTTTTGTGCTAAGTTAACTCTCATTAAACACCAGAGAGCACACG 1920

Qy   1921 caggggggaagcctcatgtgtgcagggagtgtgggcaaggctttagccggcagtcacacc 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 CAGGGGGGAAGCCTCATGTGTGCAGGGAGTGTGGGCAAGGCTTTAGCCGGCAGTCACACC 1980

Qy   1981 tcattagacaccagaggacacattcaggagagaagccttatatttgcagaaagtgtggac 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 TCATTAGACACCAGAGGACACATTCAGGAGAGAAGCCTTATATTTGCAGAAAGTGTGGAC 2040

Qy   2041 ggggctttagtcggaagtccaaccttatcagacatcagaggacacactcaggatagaaac 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 GGGGCTTTAGTCGGAAGTCCAACCTTATCAGACATCAGAGGACACACTCAGGATAGAAAC 2100

Qy   2101 tttatgtgtatagggaatgtggtacagcctttagccaggagtcatacttcatcagacacc 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 TTTATGTGTATAGGGAATGTGGTACAGCCTTTAGCCAGGAGTCATACTTCATCAGACACC 2160

Qy   2161 agaggacacacacagtgctgtggctttttcagccattgctagataccaaagtggagacat 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 AGAGGACACACACAGTGCTGTGGCTTTTTCAGCCATTGCTAGATACCAAAGTGGAGACAT 2220

Qy   2221 tctgtgtgtgattatgcatgagactgtactggtaagacttgtatctccatccacctgaag 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 TCTGTGTGTGATTATGCATGAGACTGTACTGGTAAGACTTGTATCTCCATCCACCTGAAG 2280

Qy   2281 gagaattgctggctcattttcaggagccctgcccttcctcactgtggatggtgggttgtg 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 GAGAATTGCTGGCTCATTTTCAGGAGCCCTGCCCTTCCTCACTGTGGATGGTGGGTTGTG 2340

Qy   2341 gaaacccggtcaggtaatgatagtggcaggaggcagtcaaatgcccaggcagataggggt 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 GAAACCCGGTCAGGTAATGATAGTGGCAGGAGGCAGTCAAATGCCCAGGCAGATAGGGGT 2400

Qy   2401 gggtacctggtgaaacccaaccttaaagctgaagacagtcccggctaaatcctcatactg 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 GGGTACCTGGTGAAACCCAACCTTAAAGCTGAAGACAGTCCCGGCTAAATCCTCATACTG 2460

Qy   2461 aattgagaacctgtcttcccatttggtgtgctttcctccgattgatcccaacccttcacc 2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461 AATTGAGAACCTGTCTTCCCATTTGGTGTGCTTTCCTCCGATTGATCCCAACCCTTCACC 2520

Qy   2521 tattttacgtatacctgccctttcctaattggtttttacactgctgtgcccacctttga 2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521 TATTTTACGTATACCTGCCCTTTCCTAATTGGTTTTTACACTGCTGTGCCCACCTTTTGA 2580

Qy   2581 gtggtgcctttgcatacttacaaatcagtcaacgtgtattcccctattctgagcccataa 2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581 GTGGTGCCTTTGCATACTTACAAATCAGTCAACGTGTATTCCCCTATTCTGAGCCCATAA 2640

Qy   2641 aagacccagactcagctgcagtgaggagagaaatcaccctgctgtggaggttggggacca 2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641 AAGACCCAGACTCAGCTGCAGTGAGGAGAGAAATCACCCTGCTGTGGAGGTTGGGGACCA 2700

Qy   2701 ctccctgcatcccctctccactgagagctgttcttttgctcaataaaattcttttctacc 2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2701 CTCCCTGCATCCCCTCTCCACTGAGAGCTGTTCTTTTGCTCAATAAAATTCTTTTCTACC 2760

Qy   2761 catcctcaccct 2772
          ||||||||||||
Db   2761 CATCCTCACCCT 2772
```

RESULT 2
AK056243
LOCUS AK056243 2854 bp mRNA linear PRI 31-OCT-2001
DEFINITION Homo sapiens cDNA FLJ31681 fis, clone NT2RI2005315, highly similar to Homo sapiens mRNA for HKR1.
ACCESSION AK056243
VERSION AK056243.1 GI:16551590
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens teratocarcinoma cell_line:NT2 cDNA to mRNA, clone_lib:NT2RI2 clone:NT2RI2005315.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Ishibashi,T., Kanehori,K., Yosida,M., Watanabe,S., Ishida,S., Ono,Y., Hotuta,T., Hiraoka,S., Murakawa,K., Takiguchi,S., Kusano,J., Watanabe,M., Fujimori,K., Tanai,H., Ishida,M., Yamashita,H., Chiba,Y., Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K., Wagatsuma,M., Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2854)
AUTHORS Isogai,T., Otsuki,T. and Sugiyama,T.
TITLE Direct Submission
JOURNAL Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3951, Fax:81-438-52-3952)
COMMENT NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: RAB and HRI.
FEATURES Location/Qualifiers
source 1. .2854
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NT2RI2005315"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_lib="NT2RI2"
/note="cloning vector: pME18SFL3-mRNA from NT2 neuronal precursor cells treated 2-weeks mitotic inhibitor after 5-weeks retinoic acid (RA) induction.-majorly NT2 neuron"
BASE COUNT 761 a 692 c 754 g 647 t
ORIGIN

Query Match 86.0%; Score 2383; DB 9; Length 2854;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 2433; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy    334 aaattccatcttctaaaccaaaactcattgctcagctggagcgaggggaagcgccctgga 393
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 AAATTCCATCTTCTAAACCAAAACTCATTGCTCAGCTGGAGCGAGGGGAAGCGCCCTGGA 480

Qy    394 gagaggagagaaaatgtccactggacctctgtccagaatcgaagccagaaattcaacttta 453
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GAGAGGAGAGAAAATGTCCACTGGACCTCTGTCCAGAATCGAAGCCAGAAATTCAACTTA 540

Qy    454 gtccctcctgccctctgattttctccagtcagcaagctctcagccaacatgtgtggctga 513
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GTCCCTCCTGCCCTCTGATTTTCTCCAGTCAGCAAGCTCTCAGCCAACATGTGTGGCTGA 600

Qy    514 gtcatctctctcagctgttttcaagtttatgggcaggaaatcctctccacctgggaaaac 573
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GTCATCTCTCTCAGCTGTTTTCAAGTTTATGGGCAGGAAATCCTCTCCACCTGGGAAAAC 660

Qy    574 actatccagaagatcagaaacaacagcaggatccattctgctttagtggcaaagcagaat 633
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 ACTATCCAGAAGATCAGAAACAACAGCAGGATCCATTCTGCTTTAGTGGCAAAGCAGAAT 720

Qy    634 ggattcaagagggagaagactccagactcctgtttgggagagtaagcaaaaatggcactt 693
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GGATTCAAGAGGGAGAAGACTCCAGACTCCTGTTTGGGAGAGTAAGCAAAAATGGCACTT 780

Qy    694 caaaggcactttccagcccacctgaagaacaacagccagcacagtccaaggaagacaaca 753
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CAAAGGCACTTTCCAGCCCACCTGAAGAACAACAGCCAGCACAGTCCAAGGAAGACAACA 840

Qy    754 cagtggtggatatagggtccagccctgaacggagggcagatctagaggaaacagacaaag 813
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CAGTGGTGGATATAGGGTCCAGCCCTGAACGGAGGGCAGATCTAGAGGAAACAGACAAAG 900

Qy    814 tattgcatggtttagaagtctcaggatttggagaaatcaaatatgaagagtttgggccag 873
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TATTGCATGGTTTAGAAGTCTCAGGATTTGGAGAAATCAAATATGAAGAGTTTGGGCCAG 960

Qy    874 gctttatcaaggagtcaaacctccttagcctccagaagacacaaactggggagacacctt 933
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GCTTTATCAAGGAGTCAAACCTCCTTAGCCTCCAGAAGACACAAACTGGGGAGACACCTT 1020

Qy    934 acatgtacactgagtggggagacagctttggcagtatgtcagtcctcatcaaaaacccaa 993
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 ACATGTACACTGAGTGGGGAGACAGCTTTGGCAGTATGTCAGTCCTCATCAAAAACCCAA 1080

Qy    994 ggacacactctgggggaaagccttatgtgtgcagggaatgtgggcgaggctttacgtgga 1053
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 GGACACACTCTGGGGGAAAGCCTTATGTGTGCAGGGAATGTGGGCGAGGCTTTACGTGGA 1140

Qy   1054 agtcaaacctgatcacacatcagaggacacactcaggggagaaaccttatgtgtgcaagg 1113
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 AGTCAAACCTGATCACACATCAGAGGACACACTCAGGGGAGAAACCTTATGTGTGCAAGG 1200

Qy   1114 attgtggacgaggctttacttggaagtcgaacctctttacacatcagcggacacactcag 1173
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 ATTGTGGACGAGGCTTTACTTGGAAGTCGAACCTCTTTACACATCAGCGGACACACTCAG 1260

Qy   1174 ggctcaagccttatgtgtgcaaggaatgtgggcagagctttagcctgaagtcaaacctca 1233
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GGCTCAAGCCTTATGTGTGCAAGGAATGTGGGCAGAGCTTTAGCCTGAAGTCAAACCTCA 1320

Qy   1234 ttacccaccagagggcgcacactggggagaagccttatgtttgcagggaatgtgggcgtg 1293
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 TTACCCACCAGAGGGCGCACACTGGGGAGAAGCCTTATGTTTGCAGGGAATGTGGGCGTG 1380

Qy   1294 gctttcgccagcattcacacctggtcagacacaagaggacacattcaggagagaagcctt 1353
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 GCTTTCGCCAGCATTCACACCTGGTCAGACACAAGAGGACACATTCAGGAGAGAAGCCTT 1440

Qy   1354 acatttgcagggagtgtgagcaaggctttagccagaagtcacacctcatcagacacttaa 1413
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 ACATTTGCAGGGAGTGTGAGCAAGGCTTTAGCCAGAAGTCACACCTCATCAGACACTTAA 1500

Qy   1414 ggacacacacaggagagaagccttatgtatgcacagaatgtgggcgtcactttagctgga 1473
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 GGACACACACAGGAGAGAAGCCTTATGTATGCACAGAATGTGGGCGTCACTTTAGCTGGA 1560

Qy   1474 aatcaaacctcaaaacacaccagaggacacactcaggggttaaaccttatgtctgcctgg 1533
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 AATCAAACCTCAAAACACACCAGAGGACACACTCAGGGGTTAAACCTTATGTCTGCCTGG 1620

Qy   1534 agtgcgggcagtgctttagcctgaagtcaaaccttaacaaacaccagaggtcacacacgg 1593
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 AGTGCGGGCAGTGCTTTAGCCTGAAGTCAAACCTTAACAAACACCAGAGGTCACACACGG 1680

Qy   1594 gggagaagccatttgtatgtacggagtgtgggcgaggctttacccggaaatcaaccctga 1653
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 GGGAGAAGCCATTTGTATGTACGGAGTGTGGGCGAGGCTTTACCCGGAAATCAACCCTGA 1740

Qy   1654 tcacgcaccagaggacacactcaggggagaagccatttgtatgtgctgagtgtggacgag 1713
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 TCACGCACCAGAGGACACACTCAGGGGAGAAGCCATTTGTATGTGCTGAGTGTGGACGAG 1800

Qy   1714 gctttaatgataagtccaccctcatttcacaccagaggacacattcaggggaaaagcctt 1773
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 GCTTTAATGATAAGTCCACCCTCATTTCACACCAGAGGACACATTCAGGGGAAAAGCCTT 1860

Qy   1774 ttatgtgcagggagtgtggcagaaggtttcggcagaagcctaacctgtttaggcacaaga 1833
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 TTATGTGCAGGGAGTGTGGCAGAAGGTTTCGGCAGAAGCCTAACCTGTTTAGGCACAAGA 1920

Qy   1834 gggcacactcaggtgcctttgtgtgcagggagtgtgggcaaggcttttgtgctaagttaa 1893
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 GGGCACACTCAGGTGCCTTTGTGTGCAGGGAGTGTGGGCAAGGCTTTTGTGCTAAGTTAA 1980

Qy   1894 ctctcattaaacaccagagagcacacgcaggggggaagcctcatgtgtgcagggagtgtg 1953
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 CTCTCATTAAACACCAGAGAGCACACGCAGGGGGGAAGCCTCATGTGTGCAGGGAGTGTG 2040

Qy   1954 ggcaaggctttagccggcagtcacacctcattagacaccagaggacacattcaggagaga 2013
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 GGCAAGGCTTTAGCCGGCAGTCACACCTCATTAGACACCAGAGGACACATTCAGGAGAGA 2100

Qy   2014 agccttatatttgcagaaagtgtggacggggctttagtcggaagtccaaccttatcagac 2073
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 AGCCTTATATTTGCAGAAAGTGTGGACGGGGCTTTAGTCGGAAGTCCAACCTTATCAGAC 2160

Qy   2074 atcagaggacacactcaggatagaaactttatgtgtatagggaatgtggtacagccttta 2133
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 ATCAGAGGACACACTCAGGATAGAAACTTTATGTGTATAGGGAATGTGGTACAGCCTTTA 2220

Qy   2134 gccaggagtcatacttcatcagacaccagaggacacacacagtgctgtggctttttcagc 2193
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 GCCAGGAGTCATACTTCATCAGACACCAGAGGACACACACAGTGCTGTGGCTTTTTCAGC 2280

Qy   2194 cattgctagataccaaagtggagacattctgtgtgtgattatgcatgagactgtactggt 2253
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 CATTGCTAGATACCAAAGTGGAGACATTCTGTGTGTGATTATGCATGAGACTGTACTGGT 2340

Qy   2254 aagacttgtatctccatccacctgaaggagaattgctggctcattttcaggagccctgcc 2313
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 AAGACTTGTATCTCCATCCACCTGAAGGAGAATTGCTGGCTCATTTTCAGGAGCCCTGCC 2400

Qy   2314 cttcctcactgtggatggtgggttgtggaaacccggtcaggtaatgatagtggcaggagg 2373
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 CTTCCTCACTGTGGATGGTGGGTTGTGGAAACCCGGTCAGGTAATGATAGTGGCAGGAGG 2460

Qy   2374 cagtcaaatgcccaggcagataggggtgggtacctggtgaaacccaaccttaaagctgaa 2433
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461 CAGTCAAATGCCCAGGCAGATAGGGGTGGGTACCTGGTGAAACCCAACCTTAAAGCTGAA 2520

Qy   2434 gacagtcccggctaaatcctcatactgaattgagaacctgtcttcccatttggtgtgctt 2493
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521 GACAGTCCCGGCTAAATCCTCATACTGAATTGAGAACCTGTCTTCCCATTTGGTGTGCTT 2580
```

```
Qy    2494 tcctccgattgatcccaacccttcacctattttacgtatacctgccctttcctaattggt 2553
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2581 TCCTCCGATTGATCCCAACCCTTCACCTATTTTACGTATACCTGCCCTTTCCTAATTGGT 2640

Qy    2554 ttttacactgctgtgcccacctttgagtggtgcctttgcatacttacaaatcagtcaac 2613
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2641 TTTTACACTGCTGTGCCCACCTTTTGAGTGGTGCCTTTGCATACTTACAAATCAGTCAAC 2700

Qy    2614 gtgtattcccctattctgagcccataaaagacccagactcagctgcagtgaggagagaaa 2673
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2701 GTGTATTCCCCTATTCTGAGCCCATAAAAGACCCAGACTCAGCTGCAGTGAGGAGAGAAA 2760

Qy    2674 tcaccctgctgtggaggttggggaccactccctgcatcccctctccactgagagctgttc 2733
           ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db    2761 TCACCCTGCTGTGGGGGTTGGGGACCACTCCCTGCATCCCCTCTCCACTGAGAGCTGTTC 2820

Qy    2734 ttttgctcaataaaattcttttctacccatcctc 2767
           |||||||||||||||||||||||||||||||||
Db    2821 TTTTGCTCAATAAAATTCTTTTCTACCCATCCTC 2854


RESULT   3
AC008806/c
LOCUS       AC008806              135173 bp    DNA     linear   PRI 27-APR-2000
DEFINITION  Homo sapiens chromosome 19 clone CTD-2086O20, complete sequence.
ACCESSION   AC008806
VERSION     AC008806.4  GI:7656695
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 135173)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 135173)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 135173)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-APR-2000) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Apr 27, 2000 this sequence version replaced gi:6600982.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 99.5% of Sequence;
            Estimated Total Number of Errors is 0.6.
            SHGC-141611 G63592
            SHGC-37349 G30531
            SHGC-86030 G61627.
FEATURES             Location/Qualifiers
     source          1. .135173
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="19"
                     /clone="CTD-2086O20"
BASE COUNT    41590 a  27234 c  27240 g  39109 t
ORIGIN


  Query Match           81.0%;  Score 2244;  DB 9;  Length 135173;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 2344;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy     427 cagaatcgaagccagaaattcaacttagtccctcctgccctctgattttctccagtcagc 486
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  126941 CAGAATCGAAGCCAGAAATTCAACTTAGTCCCTCCTGCCCTCTGATTTTCTCCAGTCAGC 126882

Qy     487 aagctctcagccaacatgtgtggctgagtcatctctctcagctgttttcaagtttatggg 546
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  126881 AAGCTCTCAGCCAACATGTGTGGCTGAGTCATCTCTCTCAGCTGTTTTCAAGTTTATGGG 126822

Qy     547 caggaaatcctctccacctgggaaaacactatccagaagatcagaaacaacagcaggatc 606
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  126821 CAGGAAATCCTCTCCACCTGGGAAAACACTATCCAGAAGATCAGAAACAACAGCAGGATC 126762

Qy     607 cattctgctttagtggcaaagcagaatggattcaagagggagaagactccagactcctgt 666
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  126761 CATTCTGCTTTAGTGGCAAAGCAGAATGGATTCAAGAGGGAGAAGACTCCAGACTCCTGT 126702

Qy     667 ttgggagagtaagcaaaaatggcacttcaaaggcactttccagcccacctgaagaacaac 726
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  126701 TTGGGAGAGTAAGCAAAAATGGCACTTCAAAGGCACTTTCCAGCCCACCTGAAGAACAAC 126642

Qy     727 agccagcacagtccaaggaagacaacacagtggtggatatagggtccagccctgaacgga 786
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  126641 AGCCAGCACAGTCCAAGGAAGACAACACAGTGGTGGATATAGGGTCCAGCCCTGAACGGA 126582

Qy     787 gggcagatctagaggaaacagacaaagtattgcatggtttagaagtctcaggatttggag 846
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  126581 GGGCAGATCTAGAGGAAACAGACAAAGTATTGCATGGTTTAGAAGTCTCAGGATTTGGAG 126522

Qy     847 aaatcaaatatgaagagtttgggccaggctttatcaaggagtcaaacctccttagcctcc 906
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  126521 AAATCAAATATGAAGAGTTTGGGCCAGGCTTTATCAAGGAGTCAAACCTCCTTAGCCTCC 126462

Qy     907 agaagacacaaactggggagacaccttacatgtacactgagtggggagacagctttggca 966
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  126461 AGAAGACACAAACTGGGGAGACACCTTACATGTACACTGAGTGGGGAGACAGCTTTGGCA 126402

Qy     967 gtatgtcagtcctcatcaaaaacccaaggacacactctgggggaaagccttatgtgtgca 1026
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  126401 GTATGTCAGTCCTCATCAAAAACCCAAGGACACACTCTGGGGGAAAGCCTTATGTGTGCA 126342

Qy    1027 gggaatgtgggcgaggctttacgtggaagtcaaacctgatcacacatcagaggacacact 1086
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  126341 GGGAATGTGGGCGAGGCTTTACGTGGAAGTCAAACCTGATCACACATCAGAGGACACACT 126282

Qy    1087 caggggagaaaccttatgtgtgcaaggattgtggacgaggctttacttggaagtcgaacc 1146
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  126281 CAGGGGAGAAACCTTATGTGTGCAAGGATTGTGGACGAGGCTTTACTTGGAAGTCGAACC 126222

Qy    1147 tctttacacatcagcggacacactcagggctcaagccttatgtgtgcaaggaatgtgggc 1206
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  126221 TCTTTACACATCAGCGGACACACTCAGGGCTCAAGCCTTATGTGTGCAAGGAATGTGGGC 126162

Qy    1207 agagctttagcctgaagtcaaacctcattacccaccagagggcgcacactggggagaagc 1266
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  126161 AGAGCTTTAGCCTGAAGTCAAACCTCATTACCCACCAGAGGGCGCACACTGGGGAGAAGC 126102

Qy    1267 cttatgtttgcagggaatgtgggcgtggctttcgccagcattcacacctggtcagacaca 1326
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  126101 CTTATGTTTGCAGGGAATGTGGGCGTGGCTTTCGCCAGCATTCACACCTGGTCAGACACA 126042

Qy    1327 agaggacacattcaggagagaagccttacatttgcagggagtgtgagcaaggctttagcc 1386
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  126041 AGAGGACACATTCAGGAGAGAAGCCTTACATTTGCAGGGAGTGTGAGCAAGGCTTTAGCC 125982

Qy    1387 agaagtcacacctcatcagacacttaaggacacacacaggagagaagccttatgtatgca 1446
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  125981 AGAAGTCACACCTCATCAGACACTTAAGGACACACACAGGAGAGAAGCCTTATGTATGCA 125922

Qy    1447 cagaatgtgggcgtcactttagctggaaatcaaacctcaaaacacaccagaggacacact 1506
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  125921 CAGAATGTGGGCGTCACTTTAGCTGGAAATCAAACCTCAAAACACACCAGAGGACACACT 125862

Qy    1507 caggggttaaaccttatgtctgcctggagtgcgggcagtgctttagcctgaagtcaaacc 1566
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  125861 CAGGGGTTAAACCTTATGTCTGCCTGGAGTGCGGGCAGTGCTTTAGCCTGAAGTCAAACC 125802
```

```
Qy   1567 ttaacaaacaccagaggtcacacacgggggagaagccatttgtatgtacggagtgtgggc 1626
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 125801 TTAACAAACACCAGAGGTCACACACGGGGGAGAAGCCATTTGTATGTACGGAGTGTGGGC 125742

Qy   1627 gaggctttacccggaaatcaaccctgatcacgcaccagaggacacactcaggggagaagc 1686
          ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db 125741 GAGGCTTTACCCGGAAATCAACCCTGAGCACGCACCAGAGGACACACTCAGGGGAGAAGC 125682

Qy   1687 catttgtatgtgctgagtgtggacgaggctttaatgataagtccaccctcatttcacacc 1746
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 125681 CATTTGTATGTGCTGAGTGTGGACGAGGCTTTAATGATAAGTCCACCCTCATTTCACACC 125622

Qy   1747 agaggacacattcaggggaaaagccttttatgtgcagggagtgtggcagaaggtttcggc 1806
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 125621 AGAGGACACATTCAGGGGAAAAGCCTTTTATGTGCAGGGAGTGTGGCAGAAGGTTTCGGC 125562

Qy   1807 agaagcctaacctgtttaggcacaagagggcacactcaggtgcctttgtgtgcagggagt 1866
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 125561 AGAAGCCTAACCTGTTTAGGCACAAGAGGGCACACTCAGGTGCCTTTGTGTGCAGGGAGT 125502

Qy   1867 gtgggcaaggcttttgtgctaagttaactctcattaaacaccagagagcacacgcagggg 1926
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 125501 GTGGGCAAGGCTTTTGTGCTAAGTTAACTCTCATTAAACACCAGAGAGCACACGCAGGGG 125442

Qy   1927 ggaagcctcatgtgtgcagggagtgtgggcaaggctttagccggcagtcacacctcatta 1986
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 125441 GGAAGCCTCATGTGTGCAGGGAGTGTGGGCAAGGCTTTAGCCGGCAGTCACACCTCATTA 125382

Qy   1987 gacaccagaggacacattcaggagagaagccttatatttgcagaaagtgtggacggggct 2046
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 125381 GACACCAGAGGACACATTCAGGAGAGAAGCCTTATATTTGCAGAAAGTGTGGACGGGGCT 125322

Qy   2047 ttagtcggaagtccaaccttatcagacatcagaggacacactcaggatagaaactttatg 2106
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 125321 TTAGTCGGAAGTCCAACCTTATCAGACATCAGAGGACACACTCAGGATAGAAACTTTATG 125262

Qy   2107 tgtatagggaatgtggtacagcctttagccaggagtcatacttcatcagacaccagagga 2166
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 125261 TGTATAGGGAATGTGGTACAGCCTTTAGCCAGGAGTCATACTTCATCAGACACCAGAGGA 125202

Qy   2167 cacacacagtgctgtggctttttcagccattgctagataccaaagtggagacattctgtg 2226
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 125201 CACACACAGTGCTGTGGCTTTTTCAGCCATTGCTAGATACCAAAGTGGAGACATTCTGTG 125142

Qy   2227 tgtgattatgcatgagactgtactggtaagacttgtatctccatccacctgaaggagaat 2286
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 125141 TGTGATTATGCATGAGACTGTACTGGTAAGACTTGTATCTCCATCCACCTGAAGGAGAAT 125082

Qy   2287 tgctggctcattttcaggagccctgcccttcctcactgtggatggtgggttgtggaaacc 2346
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 125081 TGCTGGCTCATTTTCAGGAGCCCTGCCCTTCCTCACTGTGGATGGTGGGTTGTGGAAACC 125022

Qy   2347 cggtcaggtaatgatagtggcaggaggcagtcaaatgcccaggcagataggggtgggtac 2406
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 125021 CGGTCAGGTAATGATAGTGGCAGGAGGCAGTCAAATGCCCAGGCAGATAGGGGTGGGTAC 124962

Qy   2407 ctggtgaaacccaaccttaaagctgaagacagtcccggctaaatcctcatactgaattga 2466
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 124961 CTGGTGAAACCCAACCTTAAAGCTGAAGACAGTCCCGGCTAAATCCTCATACTGAATTGA 124902

Qy   2467 gaacctgtcttcccatttggtgtgctttcctccgattgatcccaacccttcacctatttt 2526
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 124901 GAACCTGTCTTCCCATTTGGTGTGCTTTCCTCCGATTGATCCCAACCCTTCACCTATTTT 124842

Qy   2527 acgtatacctgccctttcctaattggtttttacactgctgtgcccacctttgagtggtg 2586
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 124841 ACGTATACCTGCCCTTTCCTAATTGGTTTTTACACTGCTGTGCCCACCTTTTGAGTGGTG 124782

Qy   2587 cctttgcatacttacaaatcagtcaacgtgtattcccctattctgagcccataaaagacc 2646
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 124781 CCTTTGCATACTTACAAATCAGTCAACGTGTATTCCCCTATTCTGAGCCCATAAAAGACC 124722

Qy   2647 cagactcagctgcagtgaggagagaaatcaccctgctgtggaggttggggaccactccct 2706
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 124721 CAGACTCAGCTGCAGTGAGGAGAGAAATCACCCTGCTGTGGGGGTTGGGGACCACTCCCT 124662

Qy   2707 gcatcccctctccactgagagctgttcttttgctcaataaaattcttttctacccatcct 2766
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 124661 GCATCCCCTCTCCACTGAGAGCTGTTCTTTTGCTCAATAAAATTCTTTTCTACCCATCCT 124602

Qy   2767 caccct 2772
          ||||||
Db 124601 CACCCT 124596
```

RESULT 4
AC016590

```
LOCUS       AC016590               253217 bp    DNA     linear   HTG 20-APR-2001
DEFINITION  Homo sapiens chromosome 19 clone CTD-3220F14, WORKING DRAFT
            SEQUENCE, 39 unordered pieces.
ACCESSION   AC016590
VERSION     AC016590.6  GI:13699590
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 253217)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Sequencing of Human Chromosome 19
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 253217)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-DEC-1999) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT     On Apr 20, 2001 this sequence version replaced gi:7711568.
            -------Genome Center
            Center: Joint Genome Institute
            Center Code: JGI
            Web site: http://www.jgi.doe.gov
            -------
            Project Information
            Center Project Name: 940643, BC905667
            Center clone name: CITB-E1_3220F14
            -------
            Summary Statistics
            Consensus quality: 207648 bases at least Q40
            Consensus quality: 225569 bases at least Q30
            Consensus quality: 232428 bases at least Q20
            Estimated insert size: 250510; agarose-fp estimation
            Estimated insert size: 249417; sum-of-contigs estimation
            Quality coverage:  9.03 in Q20 bases; agarose-fp estimation
            Quality coverage:  9.07 in Q20 bases; sum-of-contigs estimation.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 39 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1019: contig of 1019 bp in length
            *     1020     1119: gap of unknown length
            *     1120     2152: contig of 1033 bp in length
            *     2153     2252: gap of unknown length
            *     2253     3701: contig of 1449 bp in length
            *     3702     3801: gap of unknown length
            *     3802     5053: contig of 1252 bp in length
            *     5054     5153: gap of unknown length
            *     5154     6184: contig of 1031 bp in length
            *     6185     6284: gap of unknown length
            *     6285     7809: contig of 1525 bp in length
            *     7810     7909: gap of unknown length
            *     7910     8986: contig of 1077 bp in length
            *     8987     9086: gap of unknown length
```

```
          *     9087    10090: contig of 1004 bp in length
          *    10091    10190: gap of unknown length
          *    10191    11225: contig of 1035 bp in length
          *    11226    11325: gap of unknown length
          *    11326    12365: contig of 1040 bp in length
          *    12366    12465: gap of unknown length
          *    12466    14535: contig of 2070 bp in length
          *    14536    14635: gap of unknown length
          *    14636    16489: contig of 1854 bp in length
          *    16490    16589: gap of unknown length
          *    16590    19109: contig of 2520 bp in length
          *    19110    19209: gap of unknown length
          *    19210    21841: contig of 2632 bp in length
          *    21842    21941: gap of unknown length
          *    21942    23735: contig of 1794 bp in length
          *    23736    23835: gap of unknown length
          *    23836    25861: contig of 2026 bp in length
          *    25862    25961: gap of unknown length
          *    25962    29117: contig of 3156 bp in length
          *    29118    29217: gap of unknown length
          *    29218    31907: contig of 2690 bp in length
          *    31908    32007: gap of unknown length
          *    32008    36364: contig of 4357 bp in length
          *    36365    36464: gap of unknown length
          *    36465    43309: contig of 6845 bp in length
          *    43310    43409: gap of unknown length
          *    43410    48185: contig of 4776 bp in length
          *    48186    48285: gap of unknown length
          *    48286    55626: contig of 7341 bp in length
          *    55627    55726: gap of unknown length
          *    55727    65781: contig of 10055 bp in length
          *    65782    65881: gap of unknown length
          *    65882    69673: contig of 3792 bp in length
          *    69674    69773: gap of unknown length
          *    69774    79389: contig of 9616 bp in length
          *    79390    79489: gap of unknown length
          *    79490    91453: contig of 11964 bp in length
          *    91454    91553: gap of unknown length
          *    91554    96671: contig of 5118 bp in length
          *    96672    96771: gap of unknown length
          *    96772   100544: contig of 3773 bp in length
          *   100545   100644: gap of unknown length
          *   100645   111423: contig of 10779 bp in length
          *   111424   111523: gap of unknown length
          *   111524   115278: contig of 3755 bp in length
          *   115279   115378: gap of unknown length
          *   115379   124011: contig of 8633 bp in length
          *   124012   124111: gap of unknown length
          *   124112   138475: contig of 14364 bp in length
          *   138476   138575: gap of unknown length
          *   138576   140613: contig of 2038 bp in length
          *   140614   140713: gap of unknown length
          *   140714   144949: contig of 4236 bp in length
          *   144950   145049: gap of unknown length
          *   145050   164352: contig of 19303 bp in length
          *   164353   164452: gap of unknown length
          *   164453   170540: contig of 6088 bp in length
          *   170541   170640: gap of unknown length
          *   170641   185046: contig of 14406 bp in length
          *   185047   185146: gap of unknown length
          *   185147   222209: contig of 37063 bp in length
          *   222210   222309: gap of unknown length
          *   222310   253217: contig of 30908 bp in length.
FEATURES             Location/Qualifiers
     source          1. .253217
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="19"
                     /clone="CTD-3220F14"
                     /clone_lib="CalTech human BAC library D"
BASE COUNT    62246 a  62867 c  62532 g  61752 t   3820 others
ORIGIN
```

```
  Query Match             81.0%;  Score 2244;  DB 2;  Length 253217;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 2344;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy     427 cagaatcgaagccagaaattcaacttagtccctcctgccctctgattttctccagtcagc 486
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   74511 CAGAATCGAAGCCAGAAATTCAACTTAGTCCCTCCTGCCCTCTGATTTTCTCCAGTCAGC 74570

Qy     487 aagctctcagccaacatgtgtggctgagtcatctctctcagctgttttcaagtttatggg 546
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   74571 AAGCTCTCAGCCAACATGTGTGGCTGAGTCATCTCTCTCAGCTGTTTTCAAGTTTATGGG 74630

Qy     547 caggaaatcctctccacctgggaaaacactatccagaagatcagaaacaacagcaggatc 606
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   74631 CAGGAAATCCTCTCCACCTGGGAAAACACTATCCAGAAGATCAGAAACAACAGCAGGATC 74690

Qy     607 cattctgctttagtggcaaagcagaatggattcaagagggagaagactccagactcctgt 666
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   74691 CATTCTGCTTTAGTGGCAAAGCAGAATGGATTCAAGAGGGAGAAGACTCCAGACTCCTGT 74750

Qy     667 ttgggagagtaagcaaaaatggcacttcaaaggcactttccagcccacctgaagaacaac 726
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   74751 TTGGGAGAGTAAGCAAAAATGGCACTTCAAAGGCACTTTCCAGCCCACCTGAAGAACAAC 74810

Qy     727 agccagcacagtccaaggaagacaacacagtggtggatatagggtccagccctgaacgga 786
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   74811 AGCCAGCACAGTCCAAGGAAGACAACACAGTGGTGGATATAGGGTCCAGCCCTGAACGGA 74870

Qy     787 gggcagatctagaggaaacagacaaagtattgcatggtttagaagtctcaggatttggag 846
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   74871 GGGCAGATCTAGAGGAAACAGACAAAGTATTGCATGGTTTAGAAGTCTCAGGATTTGGAG 74930

Qy     847 aaatcaaatatgaagagtttgggccaggctttatcaaggagtcaaacctccttagcctcc 906
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   74931 AAATCAAATATGAAGAGTTTGGGCCAGGCTTTATCAAGGAGTCAAACCTCCTTAGCCTCC 74990

Qy     907 agaagacacaaactggggagacaccttacatgtacactgagtggggagacagctttggca 966
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   74991 AGAAGACACAAACTGGGGAGACACCTTACATGTACACTGAGTGGGGAGACAGCTTTGGCA 75050

Qy     967 gtatgtcagtcctcatcaaaaacccaaggacacactctgggggaaagccttatgtgtgca 1026
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   75051 GTATGTCAGTCCTCATCAAAAACCCAAGGACACACTCTGGGGGAAAGCCTTATGTGTGCA 75110

Qy    1027 gggaatgtgggcgaggctttacgtggaagtcaaacctgatcacacatcagaggacacact 1086
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   75111 GGGAATGTGGGCGAGGCTTTACGTGGAAGTCAAACCTGATCACACATCAGAGGACACACT 75170

Qy    1087 caggggagaaaccttatgtgtgcaaggattgtggacgaggctttacttggaagtcgaacc 1146
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   75171 CAGGGGAGAAACCTTATGTGTGCAAGGATTGTGGACGAGGCTTTACTTGGAAGTCGAACC 75230

Qy    1147 tctttacacatcagcggacacactcagggctcaagccttatgtgtgcaaggaatgtgggc 1206
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   75231 TCTTTACACATCAGCGGACACACTCAGGGCTCAAGCCTTATGTGTGCAAGGAATGTGGGC 75290

Qy    1207 agagctttagcctgaagtcaaacctcattacccaccagagggcgcacactggggagaagc 1266
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   75291 AGAGCTTTAGCCTGAAGTCAAACCTCATTACCCACCAGAGGGCGCACACTGGGGAGAAGC 75350

Qy    1267 cttatgtttgcagggaatgtgggcgtggctttcgccagcattcacacctggtcagacaca 1326
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   75351 CTTATGTTTGCAGGGAATGTGGGCGTGGCTTTCGCCAGCATTCACACCTGGTCAGACACA 75410

Qy    1327 agaggacacattcaggagagaagccttacatttgcagggagtgtgagcaaggctttagcc 1386
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   75411 AGAGGACACATTCAGGAGAGAAGCCTTACATTTGCAGGGAGTGTGAGCAAGGCTTTAGCC 75470

Qy    1387 agaagtcacacctcatcagacacttaaggacacacacaggagagaagccttatgtatgca 1446
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   75471 AGAAGTCACACCTCATCAGACACTTAAGGACACACACAGGAGAGAAGCCTTATGTATGCA 75530
```

```
Qy    1447 cagaatgtgggcgtcactttagctggaaatcaaacctcaaaacacaccagaggacacact 1506
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   75531 CAGAATGTGGGCGTCACTTTAGCTGGAAATCAAACCTCAAAACACACCAGAGGACACACT 75590

Qy    1507 caggggttaaaccttatgtctgcctggagtgcgggcagtgctttagcctgaagtcaaacc 1566
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   75591 CAGGGGTTAAACCTTATGTCTGCCTGGAGTGCGGGCAGTGCTTTAGCCTGAAGTCAAACC 75650

Qy    1567 ttaacaaacaccagaggtcacacacgggggagaagccatttgtatgtacggagtgtgggc 1626
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   75651 TTAACAAACACCAGAGGTCACACACGGGGGAGAAGCCATTTGTATGTACGGAGTGTGGGC 75710

Qy    1627 gaggctttacccggaaatcaaccctgatcacgcaccagaggacacactcaggggagaagc 1686
           ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db   75711 GAGGCTTTACCCGGAAATCAACCCTGAGCACGCACCAGAGGACACACTCAGGGGAGAAGC 75770

Qy    1687 catttgtatgtgctgagtgtggacgaggctttaatgataagtccaccctcatttcacacc 1746
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   75771 CATTTGTATGTGCTGAGTGTGGACGAGGCTTTAATGATAAGTCCACCCTCATTTCACACC 75830

Qy    1747 agaggacacattcaggggaaaagccttttatgtgcagggagtgtggcagaaggtttcggc 1806
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   75831 AGAGGACACATTCAGGGGAAAAGCCTTTTATGTGCAGGGAGTGTGGCAGAAGGTTTCGGC 75890

Qy    1807 agaagcctaacctgtttaggcacaagagggcacactcaggtgcctttgtgtgcagggagt 1866
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   75891 AGAAGCCTAACCTGTTTAGGCACAAGAGGGCACACTCAGGTGCCTTTGTGTGCAGGGAGT 75950

Qy    1867 gtgggcaaggcttttgtgctaagttaactctcattaaacaccagagagcacacgcagggg 1926
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   75951 GTGGGCAAGGCTTTTGTGCTAAGTTAACTCTCATTAAACACCAGAGAGCACACGCAGGGG 76010

Qy    1927 ggaagcctcatgtgtgcagggagtgtgggcaaggctttagccggcagtcacacctcatta 1986
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   76011 GGAAGCCTCATGTGTGCAGGGAGTGTGGGCAAGGCTTTAGCCGGCAGTCACACCTCATTA 76070

Qy    1987 gacaccagaggacacattcaggagagaagccttatatttgcagaaagtgtggacggggct 2046
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   76071 GACACCAGAGGACACATTCAGGAGAGAAGCCTTATATTTGCAGAAAGTGTGGACGGGGCT 76130

Qy    2047 ttagtcggaagtccaaccttatcagacatcagaggacacactcaggatagaaactttatg 2106
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   76131 TTAGTCGGAAGTCCAACCTTATCAGACATCAGAGGACACACTCAGGATAGAAACTTTATG 76190

Qy    2107 tgtatagggaatgtggtacagcctttagccaggagtcatacttcatcagacaccagagga 2166
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   76191 TGTATAGGGAATGTGGTACAGCCTTTAGCCAGGAGTCATACTTCATCAGACACCAGAGGA 76250

Qy    2167 cacacacagtgctgtggctttttcagccattgctagataccaaagtggagacattctgtg 2226
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   76251 CACACACAGTGCTGTGGCTTTTTCAGCCATTGCTAGATACCAAAGTGGAGACATTCTGTG 76310

Qy    2227 tgtgattatgcatgagactgtactggtaagacttgtatctccatccacctgaaggagaat 2286
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   76311 TGTGATTATGCATGAGACTGTACTGGTAAGACTTGTATCTCCATCCACCTGAAGGAGAAT 76370

Qy    2287 tgctggctcattttcaggagccctgcccttcctcactgtggatggtgggttgtggaaacc 2346
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   76371 TGCTGGCTCATTTTCAGGAGCCCTGCCCTTCCTCACTGTGGATGGTGGGTTGTGGAAACC 76430

Qy    2347 cggtcaggtaatgatagtggcaggaggcagtcaaatgcccaggcagatagggggtgggtac 2406
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   76431 CGGTCAGGTAATGATAGTGGCAGGAGGCAGTCAAATGCCCAGGCAGATAGGGGTGGGTAC 76490

Qy    2407 ctggtgaaacccaaccttaaagctgaagacagtcccggctaaatcctcatactgaattga 2466
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   76491 CTGGTGAAACCCAACCTTAAAGCTGAAGACAGTCCCGGCTAAATCCTCATACTGAATTGA 76550

Qy    2467 gaacctgtcttcccatttggtgtgctttcctccgattgatcccaacccttcacctatttt 2526
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   76551 GAACCTGTCTTCCCATTTGGTGTGCTTTCCTCCGATTGATCCCAACCCTTCACCTATTTT 76610

Qy    2527 acgtatacctgccctttcctaattggtttttacactgctgtgcccacctttgagtggtg 2586
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   76611 ACGTATACCTGCCCTTTCCTAATTGGTTTTTACACTGCTGTGCCCACCTTTGAGTGGTG 76670

Qy    2587 cctttgcatacttacaaatcagtcaacgtgtattcccctattctgagcccataaaagacc 2646
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   76671 CCTTTGCATACTTACAAATCAGTCAACGTGTATTCCCCTATTCTGAGCCCATAAAAGACC 76730

Qy    2647 cagactcagctgcagtgaggagagaaatcaccctgctgtggaggttggggaccactccct 2706
           ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db   76731 CAGACTCAGCTGCAGTGAGGAGAGAAATCACCCTGCTGTGGGGGTTGGGGACCACTCCCT 76790

Qy    2707 gcatcccctctccactgagagctgttcttttgctcaataaaattcttttctacccatcct 2766
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   76791 GCATCCCCTCTCCACTGAGAGCTGTTCTTTTGCTCAATAAAATTCTTTTCTACCCATCCT 76850

Qy    2767 caccct 2772
           ||||||
Db   76851 CACCCT 76856
```

```
RESULT   5
AC093235
LOCUS       AC093235               38173 bp    DNA     linear   PRI 29-SEP-2001
DEFINITION  Homo sapiens chromosome 19 clone LLNLR-267A4, complete sequence.
ACCESSION   AC093235
VERSION     AC093235.2  GI:15808549
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 38173)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 38173)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-AUG-2001) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 38173)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-SEP-2001) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Sep 29, 2001 this sequence version replaced gi:15193369.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 99.9% of Sequence;
            Estimated Total Number of Errors is 0.
FEATURES             Location/Qualifiers
     source          1. .38173
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="19"
                     /clone="LLNLR-267A4"
BASE COUNT     9495 a    8465 c    8118 g   12095 t
ORIGIN
```

```
  Query Match             79.1%;  Score 2193;  DB 9;  Length 38173;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 2343;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy     427 cagaatcgaagccagaaattcaacttagtccctcctgcccctctgattttctccagtcagc 486
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   28888 CAGAATCGAAGCCAGAAATTCAACTTAGTCCCTCCTGCCCTCTGATTTTCTCCAGTCAGC 28947

Qy     487 aagctctcagccaacatgtgtggctgagtcatctctctcagctgttttcaagtttatggg 546
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   28948 AAGCTCTCAGCCAACATGTGTGGCTGAGTCATCTCTCTCAGCTGTTTTCAAGTTTATGGG 29007
```

```
Qy    547 caggaaatcctctccacctgggaaaacactatccagaagatcagaaacaacagcaggatc 606
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29008 CAGGAAATCCTCTCCACCTGGGAAAACACTATCCAGAAGATCAGAAACAACAGCAGGATC 29067

Qy    607 cattctgctttagtggcaaagcagaatggattcaagagggagaagactccagactcctgt 666
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29068 CATTCTGCTTTAGTGGCAAAGCAGAATGGATTCAAGAGGGAGAAGACTCCAGACTCCTGT 29127

Qy    667 ttgggagagtaagcaaaaatggcacttcaaaggcactttccagcccacctgaagaacaac 726
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29128 TTGGGAGAGTAAGCAAAAATGGCACTTCAAAGGCACTTTCCAGCCCACCTGAAGAACAAC 29187

Qy    727 agccagcacagtccaaggaagacaacacagtggtggatatagggtccagccctgaacgga 786
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29188 AGCCAGCACAGTCCAAGGAAGACAACACAGTGGTGGATATAGGGTCCAGCCCTGAACGGA 29247

Qy    787 gggcagatctagaggaaacagacaaagtattgcatggtttagaagtctcaggatttggag 846
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29248 GGGCAGATCTAGAGGAAACAGACAAAGTATTGCATGGTTTAGAAGTCTCAGGATTTGGAG 29307

Qy    847 aaatcaaatatgaagagtttgggccaggctttatcaaggagtcaaacctccttagcctcc 906
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29308 AAATCAAATATGAAGAGTTTGGGCCAGGCTTTATCAAGGAGTCAAACCTCCTTAGCCTCC 29367

Qy    907 agaagacacaaactggggagacaccttacatgtacactgagtggggagacagctttggca 966
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29368 AGAAGACACAAACTGGGGAGACACCTTACATGTACACTGAGTGGGGAGACAGCTTTGGCA 29427

Qy    967 gtatgtcagtcctcatcaaaaacccaaggacacactctgggggaaagccttatgtgtgca 1026
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29428 GTATGTCAGTCCTCATCAAAAACCCAAGGACACACTCTGGGGGAAAGCCTTATGTGTGCA 29487

Qy   1027 gggaatgtgggcgaggctttacgtggaagtcaaacctgatcacacatcagaggacacact 1086
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29488 GGGAATGTGGGCGAGGCTTTACGTGGAAGTCAAACCTGATCACACATCAGAGGACACACT 29547

Qy   1087 caggggagaaaccttatgtgtgcaaggattgtggacgaggctttacttggaagtcgaacc 1146
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29548 CAGGGGAGAAACCTTATGTGTGCAAGGATTGTGGACGAGGCTTTACTTGGAAGTCGAACC 29607

Qy   1147 tctttacacatcagcggacacactcagggctcaagccttatgtgtgcaaggaatgtgggc 1206
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29608 TCTTTACACATCAGCGGACACACTCAGGGCTCAAGCCTTATGTGTGCAAGGAATGTGGGC 29667

Qy   1207 agagctttagcctgaagtcaaacctcattacccaccagagggcgcacactggggagaagc 1266
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29668 AGAGCTTTAGCCTGAAGTCAAACCTCATTACCCACCAGAGGGCGCACACTGGGGAGAAGC 29727

Qy   1267 cttatgtttgcagggaatgtgggcgtggctttcgccagcattcacacctggtcagacaca 1326
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29728 CTTATGTTTGCAGGGAATGTGGGCGTGGCTTTCGCCAGCATTCACACCTGGTCAGACACA 29787

Qy   1327 agaggacacattcaggagagaagccttacatttgcagggagtgtgagcaaggctttagcc 1386
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29788 AGAGGACACATTCAGGAGAGAAGCCTTACATTTGCAGGGAGTGTGAGCAAGGCTTTAGCC 29847

Qy   1387 agaagtcacacctcatcagacacttaaggacacacacaggagagaagccttatgtatgca 1446
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29848 AGAAGTCACACCTCATCAGACACTTAAGGACACACACAGGAGAGAAGCCTTATGTATGCA 29907

Qy   1447 cagaatgtgggcgtcactttagctggaaatcaaacctcaaaacacaccagaggacacact 1506
          ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29908 CAGAATGTGGGCATCACTTTAGCTGGAAATCAAACCTCAAAACACACCAGAGGACACACT 29967

Qy   1507 caggggttaaaccttatgtctgcctggagtgcgggcagtgctttagcctgaagtcaaacc 1566
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29968 CAGGGGTTAAACCTTATGTCTGCCTGGAGTGCGGGCAGTGCTTTAGCCTGAAGTCAAACC 30027

Qy   1567 ttaacaaacaccagaggtcacacacgggggagaagccatttgtatgtacggagtgtgggc 1626
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  30028 TTAACAAACACCAGAGGTCACACACGGGGGAGAAGCCATTTGTATGTACGGAGTGTGGGC 30087

Qy   1627 gaggctttacccggaaatcaaccctgatcacgcaccagaggacacactcaggggagaagc 1686
          ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db  30088 GAGGCTTTACCCGGAAATCAACCCTGAGCACGCACCAGAGGACACACTCAGGGGAGAAGC 30147

Qy   1687 catttgtatgtgctgagtgtggacgaggctttaatgataagtccaccctcatttcacacc 1746
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  30148 CATTTGTATGTGCTGAGTGTGGACGAGGCTTTAATGATAAGTCCACCCTCATTTCACACC 30207

Qy   1747 agaggacacattcaggggaaaagccttttatgtgcagggagtgtggcagaaggtttcggc 1806
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  30208 AGAGGACACATTCAGGGGAAAAGCCTTTTATGTGCAGGGAGTGTGGCAGAAGGTTTCGGC 30267

Qy   1807 agaagcctaacctgtttaggcacaagagggcacactcaggtgcctttgtgtgcagggagt 1866
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  30268 AGAAGCCTAACCTGTTTAGGCACAAGAGGGCACACTCAGGTGCCTTTGTGTGCAGGGAGT 30327

Qy   1867 gtgggcaaggcttttgtgctaagttaactctcattaaacaccagagagcacacgcagggg 1926
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  30328 GTGGGCAAGGCTTTTGTGCTAAGTTAACTCTCATTAAACACCAGAGAGCACACGCAGGGG 30387

Qy   1927 ggaagcctcatgtgtgcagggagtgtgggcaaggctttagccggcagtcacacctcatta 1986
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  30388 GGAAGCCTCATGTGTGCAGGGAGTGTGGGCAAGGCTTTAGCCGGCAGTCACACCTCATTA 30447

Qy   1987 gacaccagaggacacattcaggagagaagccttatatttgcagaaagtgtggacggggct 2046
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  30448 GACACCAGAGGACACATTCAGGAGAGAAGCCTTATATTTGCAGAAAGTGTGGACGGGGCT 30507

Qy   2047 ttagtcggaagtccaaccttatcagacatcagaggacacactcaggatagaaacttttatg 2106
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  30508 TTAGTCGGAAGTCCAACCTTATCAGACATCAGAGGACACACTCAGGATAGAAACTTTATG 30567

Qy   2107 tgtatagggaatgtggtacagcctttagccaggagtcatacttcatcagacaccagagga 2166
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  30568 TGTATAGGGAATGTGGTACAGCCTTTAGCCAGGAGTCATACTTCATCAGACACCAGAGGA 30627

Qy   2167 cacacacagtgctgtggctttttcagccattgctagataccaaagtggagacattctgtg 2226
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  30628 CACACACAGTGCTGTGGCTTTTTCAGCCATTGCTAGATACCAAAGTGGAGACATTCTGTG 30687

Qy   2227 tgtgattatgcatgagactgtactggtaagacttgtatctccatccacctgaaggagaat 2286
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  30688 TGTGATTATGCATGAGACTGTACTGGTAAGACTTGTATCTCCATCCACCTGAAGGAGAAT 30747

Qy   2287 tgctggctcattttcaggagccctgcccttcctcactgtggatggtgggttgtggaaacc 2346
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  30748 TGCTGGCTCATTTTCAGGAGCCCTGCCCTTCCTCACTGTGGATGGTGGGTTGTGGAAACC 30807

Qy   2347 cggtcaggtaatgatagtggcaggaggcagtcaaatgcccaggcagataggggtgggtac 2406
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  30808 CGGTCAGGTAATGATAGTGGCAGGAGGCAGTCAAATGCCCAGGCAGATAGGGGTGGGTAC 30867

Qy   2407 ctggtgaaacccaaccttaaagctgaagacagtcccggctaaatcctcatactgaattga 2466
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  30868 CTGGTGAAACCCAACCTTAAAGCTGAAGACAGTCCCGGCTAAATCCTCATACTGAATTGA 30927

Qy   2467 gaacctgtcttcccatttggtgtgctttcctccgattgatcccaacccttcacctatttt 2526
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  30928 GAACCTGTCTTCCCATTTGGTGTGCTTTCCTCCGATTGATCCCAACCCTTCACCTATTTT 30987

Qy   2527 acgtatacctgccctttcctaattggtttttacactgctgtgcccaccttttgagtggtg 2586
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  30988 ACGTATACCTGCCCTTTCCTAATTGGTTTTTACACTGCTGTGCCCACCTTTTGAGTGGTG 31047

Qy   2587 cctttgcatacttacaaatcagtcaacgtgtattcccctattctgagcccataaaagacc 2646
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  31048 CCTTTGCATACTTACAAATCAGTCAACGTGTATTCCCCTATTCTGAGCCCATAAAAGACC 31107

Qy   2647 cagactcagctgcagtgaggagagaaatcaccctgctgtggaggttggggaccactccct 2706
          ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db  31108 CAGACTCAGCTGCAGTGAGGAGAGAAATCACCCTGCTGTGGGGGTTGGGGACCACTCCCT 31167

Qy   2707 gcatcccctctccactgagagctgttcttttgctcaataaaattcttttctacccatcct 2766
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  31168 GCATCCCCTCTCCACTGAGAGCTGTTCTTTTGCTCAATAAAATTCTTTTCTACCCATCCT 31227

Qy   2767 caccct 2772
          ||||||
Db  31228 CACCCT 31233


RESULT   6
BC004513
LOCUS       BC004513               1998 bp    mRNA    linear   PRI 12-JUL-2001
DEFINITION  Homo sapiens, clone IMAGE:3928207, mRNA, partial cds.
ACCESSION   BC004513
VERSION     BC004513.1  GI:13325426
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1998)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-MAR-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Institute for Systems Biology
            http://www.systemsbiology.org
            contact: amadan@systemsbiology.org
            Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia
            Greene, Mark Ketteman and Anuradha Madan

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAL Plate: 14 Row: b Column: 10
            This clone was selected for full length sequencing because it
            passed the following selection criteria: Hexamer frequency ORF
            analysis.
FEATURES             Location/Qualifiers
     source          1. .1998
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3928207"
                     /tissue_type="Ovary, adenocarcinoma"
                     /clone_lib="NIH_MGC_9"
                     /lab_host="DH10B-R"
                     /note="Vector: pOTB7"
     CDS             <1. .1280
                     /codon_start=3
                     /product="Unknown (protein for IMAGE:3928207)"
                     /protein_id="AAH04513.1"
                     /db_xref="GI:13325427"
                     /translation="HEVEVSGFGEIKYEEFGPGFIKESNLLSLQKTQTGETPYMYTEW
                     GDSFGSMSVLIKNPRTHSGGKPYVCRECGRGFTWKSNLITHQRTHSGEKPYVCKDCGR
                     GFTWKSNLFTHQRTHSGLKPYVCKECGQSFSLKSNLITHQRAHTGEKPYVCRECGRGF
                     RQHSHLVRHKRTHSGEKPYICRECEQGFSQKSHLIRHLRTHTGEKPYVCTECGRHFSW
                     KSNLKTHQRTHSGVKPYVCLECGQCFSLKSNLNKHQRSHTGEKPFVCTECGRGFTRKS
                     TLITHQRTHSGEKPFVCAECGRGFNDKSTLISHQRTHSGEKPFMCRECGRRFRQKPNL
                     FRHKRAHSGAFVCRECGQGFCAKLTLIKHQRAHAGGKPHVCRECGQGFSRQSHLIRHQ
                     RTHSGEKPYICRKCGRGFSRKSNLIRHQRTHSG"
BASE COUNT      575 a    458 c    510 g    455 t
ORIGIN


  Query Match           63.9%;  Score 1771;  DB 9;  Length 1998;
  Best Local Similarity  99.9%;  Pred. No. 0;
  Matches 1941;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    1;
```

```
Qy    826 tagaagtctcaggatttggagaaatcaaatatgaagagtttgggccaggctttatcaagg 885
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     10 TAGAAGTCTCAGGATTTGGAGAAATCAAATATGAAGAGTTTGGGCCAGGCTTTATCAAGG 69

Qy    886 agtcaaacctccttagcctccagaagacacaaactggggagacaccttacatgtacactg 945
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     70 AGTCAAACCTCCTTAGCCTCCAGAAGACACAAACTGGGGAGACACCTTACATGTACACTG 129

Qy    946 agtggggagacagctttggcagtatgtcagtcctcatcaaaaacccaaggacacactctg 1005
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    130 AGTGGGGAGACAGCTTTGGCAGTATGTCAGTCCTCATCAAAAACCCAAGGACACACTCTG 189

Qy   1006 ggggaaagccttatgtgtgcagggaatgtgggcgaggctttacgtggaagtcaaacctga 1065
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    190 GGGGAAAGCCTTATGTGTGCAGGGAATGTGGGCGAGGCTTTACGTGGAAGTCAAACCTGA 249

Qy   1066 tcacacatcagaggacacactcaggggagaaaccttatgtgtgcaaggattgtggacgag 1125
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    250 TCACACATCAGAGGACACACTCAGGGGAGAAACCTTATGTGTGCAAGGATTGTGGACGAG 309

Qy   1126 gctttacttggaagtcgaacctctttacacatcagcggacacactcagggctcaagcctt 1185
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    310 GCTTTACTTGGAAGTCGAACCTCTTTACACATCAGCGGACACACTCAGGGCTCAAGCCTT 369

Qy   1186 atgtgtgcaaggaatgtgggcagagctttagcctgaagtcaaacctcattacccaccaga 1245
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    370 ATGTGTGCAAGGAATGTGGGCAGAGCTTTAGCCTGAAGTCAAACCTCATTACCCACCAGA 429

Qy   1246 gggcgcacactggggagaagccttatgtttgcagggaatgtgggcgtggctttcgccagc 1305
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    430 GGGCGCACACTGGGGAGAAGCCTTATGTTTGCAGGGAATGTGGGCGTGGCTTTCGCCAGC 489

Qy   1306 attcacacctggtcagacacaagaggacacattcaggagagaagccttacatttgcaggg 1365
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    490 ATTCACACCTGGTCAGACACAAGAGGACACATTCAGGAGAGAAGCCTTACATTTGCAGGG 549

Qy   1366 agtgtgagcaaggctttagccagaagtcacacctcatcagacacttaaggacacacacag 1425
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    550 AGTGTGAGCAAGGCTTTAGCCAGAAGTCACACCTCATCAGACACTTAAGGACACACACAG 609

Qy   1426 gagagaagccttatgtatgcacagaatgtgggcgtcactttagctggaaatcaaacctca 1485
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    610 GAGAGAAGCCTTATGTATGCACAGAATGTGGGCGTCACTTTAGCTGGAAATCAAACCTCA 669

Qy   1486 aaacacaccagaggacacactcaggggttaaaccttatgtctgcctggagtgcgggcagt 1545
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    670 AAACACACCAGAGGACACACTCAGGGGTTAAACCTTATGTCTGCCTGGAGTGCGGGCAGT 729

Qy   1546 gctttagcctgaagtcaaaccttaacaaacaccagaggtcacacacgggggagaagccat 1605
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    730 GCTTTAGCCTGAAGTCAAACCTTAACAAACACCAGAGGTCACACACGGGGGAGAAGCCAT 789

Qy   1606 ttgtatgtacggagtgtgggcgaggctttacccggaaatcaaccctgatcacgcaccaga 1665
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    790 TTGTATGTACGGAGTGTGGGCGAGGCTTTACCCGGAAATCAACCCTGATCACGCACCAGA 849

Qy   1666 ggacacactcaggggagaagccatttgtatgtgctgagtgtggacgaggctttaatgata 1725
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    850 GGACACACTCAGGGGAGAAGCCATTTGTATGTGCTGAGTGTGGACGAGGCTTTAATGATA 909

Qy   1726 agtccaccctcatttcacaccagaggacacattcaggggaaaagccttttatgtgcaggg 1785
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    910 AGTCCACCCTCATTTCACACCAGAGGACACATTCAGGGGAAAAGCCTTTTATGTGCAGGG 969

Qy   1786 agtgtggcagaaggtttcggcagaagcctaacctgtttaggcacaagagggcacactcag 1845
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    970 AGTGTGGCAGAAGGTTTCGGCAGAAGCCTAACCTGTTTAGGCACAAGAGGGCACACTCAG 1029

Qy   1846 gtgcctttgtgtgcagggagtgtgggcaaggcttttgtgctaagttaactctcattaaac 1905
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1030 GTGCCTTTGTGTGCAGGGAGTGTGGGCAAGGCTTTTGTGCTAAGTTAACTCTCATTAAAC 1089
```

```
Qy   1906 accagagagcacacgcagggggaagcctcatgtgtgcagggagtgtgggcaaggcttta 1965
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1090 ACCAGAGAGCACACGCAGGGGGGAAGCCTCATGTGTGCAGGGAGTGTGGGCAAGGCTTTA 1149

Qy   1966 gccggcagtcacacctcattagacaccagaggacacattcaggagagaagccttatattt 2025
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1150 GCCGGCAGTCACACCTCATTAGACACCAGAGGACACATTCAGGAGAGAAGCCTTATATTT 1209

Qy   2026 gcagaaagtgtggacggggctttagtcggaagtccaaccttatcagacatcagaggacac 2085
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1210 GCAGAAAGTGTGGACGGGGCTTTAGTCGGAAGTCCAACCTTATCAGACATCAGAGGACAC 1269

Qy   2086 actcaggatagaaactttatgtgtatagggaatgtggtacagcctttagccaggagtcat 2145
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1270 ACTCAGGATAGAAACTTTATGTGTATAGGGAATGTGGTACAGCCTTTAGCCAGGAGTCAT 1329

Qy   2146 acttcatcagacaccagaggacacacacagtgctgtggctttttcagccattgctagata 2205
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1330 ACTTCATCAGACACCAGAGGACACACACAGTGCTGTGGCTTTTTCAGCCATTGCTAGATA 1389

Qy   2206 ccaaagtggagacattctgtgtgtgattatgcatgagactgtactggtaagacttgtatc 2265
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1390 CCAAAGTGGAGACATTCTGTGTGTGATTATGCATGAGACTGTACTGGTAAGACTTGTATC 1449

Qy   2266 tccatccacctgaaggagaattgctggctcattttcaggagccctgcccttcctcactgt 2325
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1450 TCCATCCACCTGAAGGAGAATTGCTGGCTCATTTTCAGGAGCCCTGCCCTTCCTCACTGT 1509

Qy   2326 ggatggtgggttgtggaaacccggtcaggtaatgatagtggcaggaggcagtcaaatgcc 2385
          |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db   1510 GGATGGTGGGTTGT-GAAACCCGGTCAGGTAATGATAGTGGCAGGAGGCAGTCAAATGCC 1568

Qy   2386 caggcagataggggtgggtacctggtgaaacccaaccttaaagctgaagacagtcccggc 2445
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1569 CAGGCAGATAGGGGTGGGTACCTGGTGAAACCCAACCTTAAAGCTGAAGACAGTCCCGGC 1628

Qy   2446 taaatcctcatactgaattgagaacctgtcttcccatttggtgtgctttcctccgattga 2505
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1629 TAAATCCTCATACTGAATTGAGAACCTGTCTTCCCATTTGGTGTGCTTTCCTCCGATTGA 1688

Qy   2506 tcccaacccttcacctattttacgtatacctgccctttcctaattggtttttacactgct 2565
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1689 TCCCAACCCTTCACCTATTTTACGTATACCTGCCCTTTCCTAATTGGTTTTTACACTGCT 1748

Qy   2566 gtgcccacctttgagtggtgcctttgcatacttacaaatcagtcaacgtgtattcccct 2625
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1749 GTGCCCACCTTTTGAGTGGTGCCTTTGCATACTTACAAATCAGTCAACGTGTATTCCCCT 1808

Qy   2626 attctgagcccataaaagacccagactcagctgcagtgaggagagaaatcaccctgctgt 2685
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1809 ATTCTGAGCCCATAAAAGACCCAGACTCAGCTGCAGTGAGGAGAGAAATCACCCTGCTGT 1868

Qy   2686 ggaggttggggaccactccctgcatcccctctccactgagagctgttcttttgctcaata 2745
          || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1869 GGGGGTTGGGGACCACTCCCTGCATCCCCTCTCCACTGAGAGCTGTTCTTTTGCTCAATA 1928

Qy   2746 aaattcttttctacccatcctca 2768
          |||||||||||||||||||||||
Db   1929 AAATTCTTTTCTACCCATCCTCA 1951
```

```
RESULT   7
HUMHKR1
LOCUS       HUMHKR1                   672 bp    DNA     linear   PRI 27-APR-1993
DEFINITION  Human Kruppel related gene, exon X, clone pHKR1RS.
ACCESSION   M20675 M19429
VERSION     M20675.1  GI:184108
KEYWORDS    Krueppel-related protein.
SOURCE      Human DNA, clone pHKR1RS.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 672)
  AUTHORS   Ruppert,J.M., Kinzler,K.W., Wong,A.J., Bigner,S.H., Kao,F.-T.,
            Law,M.L., Seuanez,H.N., O'Brien,S.J. and Vogelstein,B.
  TITLE     The GLI-Kruppel family of human genes
  JOURNAL   Mol. Cell. Biol. 8, 3104-3113 (1988)
  MEDLINE   89096896
COMMENT     computer-readable sequence for [1] kindly provided by J.M.Ruppert,
            20-JUN-1988.
FEATURES             Location/Qualifiers
     source          1. .672
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     exon            <1. .>672
                     /note="exon X"
BASE COUNT      195 a    155 c    184 g    138 t
ORIGIN      Chromosome 19.
```

```
  Query Match             24.2%;  Score 672;  DB 9;  Length 672;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  672;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    942 actgagtggggagacagctttggcagtatgtcagtcctcatcaaaaacccaaggacacac 1001
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ACTGAGTGGGGAGACAGCTTTGGCAGTATGTCAGTCCTCATCAAAAACCCAAGGACACAC 60

Qy   1002 tctgggggaaagccttatgtgtgcagggaatgtgggcgaggctttacgtggaagtcaaac 1061
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TCTGGGGGAAAGCCTTATGTGTGCAGGGAATGTGGGCGAGGCTTTACGTGGAAGTCAAAC 120

Qy   1062 ctgatcacacatcagaggacacactcaggggagaaaccttatgtgtgcaaggattgtgga 1121
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CTGATCACACATCAGAGGACACACTCAGGGGAGAAACCTTATGTGTGCAAGGATTGTGGA 180

Qy   1122 cgaggctttacttggaagtcgaacctctttacacatcagcggacacactcagggctcaag 1181
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CGAGGCTTTACTTGGAAGTCGAACCTCTTTACACATCAGCGGACACACTCAGGGCTCAAG 240

Qy   1182 ccttatgtgtgcaaggaatgtgggcagagctttagcctgaagtcaaacctcattacccac 1241
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CCTTATGTGTGCAAGGAATGTGGGCAGAGCTTTAGCCTGAAGTCAAACCTCATTACCCAC 300

Qy   1242 cagagggcgcacactggggagaagccttatgtttgcagggaatgtgggcgtggctttcgc 1301
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CAGAGGGCGCACACTGGGGAGAAGCCTTATGTTTGCAGGGAATGTGGGCGTGGCTTTCGC 360

Qy   1302 cagcattcacacctggtcagacacaagaggacacattcaggagagaagccttacatttgc 1361
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CAGCATTCACACCTGGTCAGACACAAGAGGACACATTCAGGAGAGAAGCCTTACATTTGC 420

Qy   1362 agggagtgtgagcaaggctttagccagaagtcacacctcatcagacacttaaggacacac 1421
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 AGGGAGTGTGAGCAAGGCTTTAGCCAGAAGTCACACCTCATCAGACACTTAAGGACACAC 480

Qy   1422 acaggagagaagccttatgtatgcacagaatgtgggcgtcactttagctggaaatcaaac 1481
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ACAGGAGAGAAGCCTTATGTATGCACAGAATGTGGGCGTCACTTTAGCTGGAAATCAAAC 540

Qy   1482 ctcaaaacacaccagaggacacactcaggggttaaaccttatgtctgcctggagtgcggg 1541
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CTCAAAACACACCAGAGGACACACTCAGGGGTTAAACCTTATGTCTGCCTGGAGTGCGGG 600

Qy   1542 cagtgctttagcctgaagtcaaaccttaacaaacaccagaggtcacacacgggggagaag 1601
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CAGTGCTTTAGCCTGAAGTCAAACCTTAACAAACACCAGAGGTCACACACGGGGGAGAAG 660

Qy   1602 ccatttgtatgt 1613
          ||||||||||||
Db    661 CCATTTGTATGT 672
```

```
RESULT   8
AX067361
LOCUS       AX067361                  678 bp    DNA     linear   PAT 24-JAN-2001
```

```
DEFINITION  Sequence 65 from Patent WO0078960.
ACCESSION   AX067361
VERSION     AX067361.1  GI:12544985
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 678)
  AUTHORS   Yuqiu,J. and Mitcham,J.L.
  TITLE     Compositions and methods for the therapy and diagnosis of breast
            cancer
  JOURNAL   Patent: WO 0078960-A 65 28-DEC-2000;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .678
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    1. .678
                     /note="n = A,T,C or G"
BASE COUNT      165 a    167 c    160 g    182 t      4 others
ORIGIN
```

Query Match 20.6%; Score 571; DB 6; Length 678;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 571; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    2078 gaggacacactcaggatagaaactttatgtgtatagggaatgtggtacagcctttagcca 2137
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9 GAGGACACACTCAGGATAGAAACTTTATGTGTATAGGGAATGTGGTACAGCCTTTAGCCA 68

Qy    2138 ggagtcatacttcatcagacaccagaggacacacacagtgctgtggctttttcagccatt 2197
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      69 GGAGTCATACTTCATCAGACACCAGAGGACACACACAGTGCTGTGGCTTTTTCAGCCATT 128

Qy    2198 gctagataccaaagtggagacattctgtgtgtgattatgcatgagactgtactggtaaga 2257
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     129 GCTAGATACCAAAGTGGAGACATTCTGTGTGTGATTATGCATGAGACTGTACTGGTAAGA 188

Qy    2258 cttgtatctccatccacctgaaggagaattgctggctcattttcaggagccctgcccttc 2317
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     189 CTTGTATCTCCATCCACCTGAAGGAGAATTGCTGGCTCATTTTCAGGAGCCCTGCCCTTC 248

Qy    2318 ctcactgtggatggtgggttgtggaaacccggtcaggtaatgatagtggcaggaggcagt 2377
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     249 CTCACTGTGGATGGTGGGTTGTGGAAACCCGGTCAGGTAATGATAGTGGCAGGAGGCAGT 308

Qy    2378 caaatgcccaggcagataggggtgggtacctggtgaaacccaaccttaaagctgaagaca 2437
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     309 CAAATGCCCAGGCAGATAGGGGTGGGTACCTGGTGAAACCCAACCTTAAAGCTGAAGACA 368

Qy    2438 gtcccggctaaatcctcatactgaattgagaacctgtcttcccatttggtgtgctttcct 2497
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     369 GTCCCGGCTAAATCCTCATACTGAATTGAGAACCTGTCTTCCCATTTGGTGTGCTTTCCT 428

Qy    2498 ccgattgatcccaacccttcacctattttacgtatacctgccctttcctaattggtttttt 2557
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     429 CCGATTGATCCCAACCCTTCACCTATTTTACGTATACCTGCCCTTTCCTAATTGGTTTTT 488

Qy    2558 acactgctgtgcccacctttgagtggtgcctttgcatacttacaaatcagtcaacgtgt 2617
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     489 ACACTGCTGTGCCCACCTTTTGAGTGGTGCCTTTGCATACTTACAAATCAGTCAACGTGT 548

Qy    2618 attcccctattctgagcccataaaagaccca 2648
           |||||||||||||||||||||||||||||||
Db     549 ATTCCCCTATTCTGAGCCCATAAAAGACCCA 579
```

```
RESULT   9
AC024911
LOCUS       AC024911              130351 bp    DNA     linear   HTG 02-MAR-2000
DEFINITION  Homo sapiens chromosome 3p clone RP11-420a23, WORKING DRAFT
            SEQUENCE, 32 unordered pieces.
ACCESSION   AC024911
VERSION     AC024911.1  GI:7143406
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 130351)
  AUTHORS   Lu,G., Zhao,Y., Lin,W., Dong,H., Wan,M., Xu,S., Gu,W., Tu,Y.,
            Jia,J., Wu,C., Zhang,C., Zhong,M., Zhou,Y., Ren,S., Fu,G., Chen,Z.
            and Huang,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-MAR-2000) Genomic Dept., Chinese National Human
            Genome Center at Shanghai, 315 Guo Shou Jing Road, Shanghai,
            Shanghai 201203, P. R. China
COMMENT     ---------------Genome Center Informations-----------------
            Center: Chinese National Human Genome Center at Shanghai
            Center Code: CHGC
            Web site: http://www.chgc.sh.cn
            Email: mhuang@chgc.sh.cn or fugang@chgc.sh.cn
             ----------------End Genome Center Informations--------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 32 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     2267: contig of 2267 bp in length
            *                   gap of unknown length
            *     2268     4378: contig of 2111 bp in length
            *                   gap of unknown length
            *     4379     6650: contig of 2272 bp in length
            *                   gap of unknown length
            *     6651     8791: contig of 2141 bp in length
            *                   gap of unknown length
            *     8792    11482: contig of 2691 bp in length
            *                   gap of unknown length
            *    11483    13872: contig of 2390 bp in length
            *                   gap of unknown length
            *    13873    16745: contig of 2873 bp in length
            *                   gap of unknown length
            *    16746    18965: contig of 2220 bp in length
            *                   gap of unknown length
            *    18966    21604: contig of 2639 bp in length
            *                   gap of unknown length
            *    21605    24066: contig of 2462 bp in length
            *                   gap of unknown length
            *    24067    26444: contig of 2378 bp in length
            *                   gap of unknown length
            *    26445    29243: contig of 2799 bp in length
            *                   gap of unknown length
            *    29244    31995: contig of 2752 bp in length
            *                   gap of unknown length
            *    31996    34538: contig of 2543 bp in length
            *                   gap of unknown length
            *    34539    37970: contig of 3432 bp in length
            *                   gap of unknown length
            *    37971    41593: contig of 3623 bp in length
            *                   gap of unknown length
            *    41594    44802: contig of 3209 bp in length
            *                   gap of unknown length
            *    44803    48691: contig of 3889 bp in length
            *                   gap of unknown length
            *    48692    52259: contig of 3568 bp in length
            *                   gap of unknown length
            *    52260    55732: contig of 3473 bp in length
            *                   gap of unknown length
            *    55733    57756: contig of 2024 bp in length
            *                   gap of unknown length
            *    57757    61389: contig of 3633 bp in length
```

```
            *                    gap of unknown length
            *    61390    65398: contig of 4009 bp in length
            *                    gap of unknown length
            *    65399    69740: contig of 4342 bp in length
            *                    gap of unknown length
            *    69741    74577: contig of 4837 bp in length
            *                    gap of unknown length
            *    74578    80165: contig of 5588 bp in length
            *                    gap of unknown length
            *    80166    86055: contig of 5890 bp in length
            *                    gap of unknown length
            *    86056    89215: contig of 3160 bp in length
            *                    gap of unknown length
            *    89216    94479: contig of 5264 bp in length
            *                    gap of unknown length
            *    94480   103136: contig of 8657 bp in length
            *                    gap of unknown length
            *   103137   113254: contig of 10118 bp in length
            *                    gap of unknown length
            *   113255   130351: contig of 17097 bp in length.
FEATURES             Location/Qualifiers
     source          1. .130351
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="3p"
                     /clone="RP11-420a23"
BASE COUNT    38785 a  26050 c  26300 g  39176 t     40 others
ORIGIN


  Query Match             1.3%;  Score 37;  DB 2;  Length 130351;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-09;
  Matches   37;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   2462 attgagaacctgtcttcccatttggtgtgctttcctc 2498
           |||||||||||||||||||||||||||||||||||
Db  69153 ATTGAGAACCTGTCTTCCCATTTGGTGTGCTTTCCTC 69189


RESULT  10
AC096898
LOCUS       AC096898              161625 bp    DNA     linear   HTG 12-JAN-2002
DEFINITION  Homo sapiens chromosome 4 clone RP11-618K19, WORKING DRAFT
            SEQUENCE, 3 unordered pieces.
ACCESSION   AC096898
VERSION     AC096898.5  GI:18139542
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 161625)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 161625)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-OCT-2001) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Jan 12, 2002 this sequence version replaced gi:17921273.

            ------------------- Genome Center --------------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu/gsc/index.shtml
            Contact: submissions@watson.wustl.edu
            ---------------- Project Information -----------------
            Center project name: H_NH0618K19
            ------------------------------------------------------
            -------------- Summary Statistics  --------------
            Sequencing vector: M13; 0%
            Sequencing vector: plasmid; 100%
            Chemistry: Dye-primer ET; 0% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 160310 bases at least Q40
            Consensus quality: 160620 bases at least Q30
            Consensus quality: 160738 bases at least Q20
            Insert size: 164000; agarose-fp
            Insert size: 161425; sum-of-contigs
            Quality coverage: 7.76 in Q20 bases; agarose-fp
            Quality coverage: 8.11 in Q20 bases; sum-of-contigs
            ------------------------------------------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 3 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     3023: contig of 3023 bp in length
            *     3024     3123: gap of unknown length
            *     3124   115853: contig of 112730 bp in length
            *   115854   115953: gap of unknown length
            *   115954   161625: contig of 45672 bp in length.
FEATURES             Location/Qualifiers
     source          1. .161625
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="4"
                     /clone="RP11-618K19"
     misc_feature    1. .3023
                     /note="assembly_name:Contig24
                     clone_end:SP6
                     vector_side:left"
     misc_feature    3124. .115853
                     /note="assembly_name:Contig26
                     clone_end:T7
                     vector_side:right"
     misc_feature    115954. .161625
                     /note="assembly_name:Contig25"
BASE COUNT    47603 a  31872 c  31765 g  50185 t    200 others
ORIGIN


  Query Match             1.3%;  Score 37;  DB 2;  Length 161625;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-09;
  Matches   37;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   2462 attgagaacctgtcttcccatttggtgtgctttcctc 2498
           |||||||||||||||||||||||||||||||||||
Db 110301 ATTGAGAACCTGTCTTCCCATTTGGTGTGCTTTCCTC 110337


RESULT  11
AC034197
LOCUS       AC034197              164264 bp    DNA     linear   HTG 29-MAY-2000
DEFINITION  Homo sapiens chromosome 3 clone RP11-728O1 map 3p, WORKING DRAFT
            SEQUENCE, 14 unordered pieces.
ACCESSION   AC034197
VERSION     AC034197.3  GI:8101273
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 164264)
  AUTHORS   Bao,J., Bao,Q., Bao,W., Bian,X., Cao,T., Chen,C., Chen,J., Ding,H.,
            Dong,W., Fan,H., Feng,X., Guan,Q., Gu,X., Guo,D., He,L., Hu,S.,
            Huang,F., Jin,Y., Kang,N., Li,C., Li,C., Li,G., Li,J., Li,L.,
            Li,S., Li,T., Liu,Y., Liu,N., Liu,B., Liu,Y., Li,W., Li,W., Li,Y.,
            Luo,J., Niu,Y., Qi,Q., Qi,X., Song,S., Sun,M., Sun,W., Sun,Y.,
            Tao,R., Wang,H., Wang,J., Wang,J., Wang,L., Wang,L., Wang,R.,
```

```
            Wang,X., Wang,X., Wang,Y., Wu,D., Wu,Q., Xie,F., Xuan,Z., Xue,Y.,
            Yan,C., Yang,X., Yu,B., Zeng,Y., Zhang,G., Zhang,H., Zhang,H.,
            Zhang,L., Zhang,M., Zhang,X., Zhang,X., Zhang,Y., Zhang,Y.,
            Zhang,Z., Zhu,B., Yu,J. and Yang,H.
  TITLE     Chromosome 3p genomic sequence
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 164264)
  AUTHORS   Zeng,Y., Hu,S., Dong,W., Wang,J., Zhang,Y., Zhang,H., Liu,B.,
            Bao,W., Sun,Y., Wu,Q., Wang,H., Yang,X., Cheng,C., Wang,Y., Niu,Y.,
            Qi,X., Li,T., Zhang,H., Liu,N., Wu,D., Yu,B., Fan,H., Liu,Y.,
            Li,G., Li,C., Bao,Q., Bao,J., Wang,X., Song,L., Zhang,L., Guo,D.,
            Huang,F., Zhang,G., Li,J., Bian,X., Zhang,M., Li,L., Feng,X., Yu,J.
            and Yang,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-APR-2000) Human Genomic Center, Institute of
            Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing
            100101, P.R.China
COMMENT     On May 29, 2000 this sequence version replaced gi:7644466.
            -------------Genome Center
            Center:Beijing Center
            Center code:Beijing
            Website:http://hgc.igtp.ac.cn
            http://www.genomics.org.cn
            Contact:hgc@igtp.ac.cn
            ------------- Project Information
            Center project name:1% project
            Center clone name: RP11-728O1
            ------------- Summary Statistics
            Sequencing vector: pUC18; 100% of reads
            Chemistry: Dye-terminator: ET 55% of reads
            Chemistry: Dye-terminator Big Dye; 45% of reads
            Assembly program: Phrap; version 0.990329
            Consensus quality: 159901 bases at least Q40
            Consensus quality: 163151 bases at least Q30
            Consensus quality: 165125 bases at least Q20
            Insert size: 158755; sum-of-contigs
            Quality coverage:   4.57x in Q20 bases;sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 14 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     2043: contig of 2043 bp in length
            *     2044     2143: gap of unknown length
            *     2144     4407: contig of 2264 bp in length
            *     4408     4507: gap of unknown length
            *     4508     5590: contig of 1083 bp in length
            *     5591     5690: gap of unknown length
            *     5691     8894: contig of 3204 bp in length
            *     8895     8994: gap of unknown length
            *     8995    12938: contig of 3944 bp in length
            *    12939    13038: gap of unknown length
            *    13039    19151: contig of 6113 bp in length
            *    19152    19251: gap of unknown length
            *    19252    24784: contig of 5533 bp in length
            *    24785    24884: gap of unknown length
            *    24885    32377: contig of 7493 bp in length
            *    32378    32477: gap of unknown length
            *    32478    41545: contig of 9068 bp in length
            *    41546    41645: gap of unknown length
            *    41646    50878: contig of 9233 bp in length
            *    50879    50978: gap of unknown length
            *    50979    62812: contig of 11834 bp in length
            *    62813    62912: gap of unknown length
            *    62913    88788: contig of 25876 bp in length
            *    88789    88888: gap of unknown length
            *    88889   115986: contig of 27098 bp in length
            *   115987   116086: gap of unknown length
            *   116087   164264: contig of 48178 bp in length.
```

```
FEATURES             Location/Qualifiers
     source          1. .164264
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="3"
                     /map="3p"
                     /clone="RP11-728O1"
     misc_feature    1. .2043
                     /note="assembly_name:Contig4"
     misc_feature    2144. .4407
                     /note="assembly_name:Contig5"
     misc_feature    4508. .5590
                     /note="assembly_name:Contig6
                     clone_end:T7
                     vector_side:right"
     misc_feature    5691. .8894
                     /note="assembly_name:Contig7"
     misc_feature    8995. .12938
                     /note="assembly_name:Contig8"
     misc_feature    13039. .19151
                     /note="assembly_name:Contig9"
     misc_feature    19252. .24784
                     /note="assembly_name:Contig10"
     misc_feature    24885. .32377
                     /note="assembly_name:Contig11"
     misc_feature    32478. .41545
                     /note="assembly_name:Contig12"
     misc_feature    41646. .50878
                     /note="assembly_name:Contig13"
     misc_feature    50979. .62812
                     /note="assembly_name:Contig14"
     misc_feature    62913. .88788
                     /note="assembly_name:Contig15"
     misc_feature    88889. .115986
                     /note="assembly_name:Contig16"
     misc_feature    116087. .164264
                     /note="assembly_name:Contig17"
BASE COUNT    46659 a  32302 c  32891 g  51080 t   1332 others
ORIGIN


  Query Match             1.3%;  Score 37;  DB 2;  Length 164264;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-09;
  Matches   37;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    2462 attgagaacctgtcttcccatttggtgtgctttcctc 2498
           |||||||||||||||||||||||||||||||||||||
Db   55166 ATTGAGAACCTGTCTTCCCATTTGGTGTGCTTTCCTC 55202


RESULT  12
AC069265
LOCUS       AC069265              165379 bp    DNA     linear   HTG 29-MAY-2000
DEFINITION  Homo sapiens chromosome 3 clone RP11-208K18 map 3p, WORKING DRAFT
            SEQUENCE, 7 unordered pieces.
ACCESSION   AC069265
VERSION     AC069265.2  GI:8101151
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 165379)
  AUTHORS   Bao,J., Bao,Q., Bao,W., Bian,X., Cao,T., Chen,C., Chen,J., Ding,H.,
            Dong,W., Fan,H., Feng,X., Guan,Q., Gu,X., Guo,D., He,L., Hu,S.,
            Huang,F., Jin,Y., Kang,N., Li,C., Li,C., Li,G., Li,J., Li,L.,
            Li,S., Li,T., Liu,Y., Liu,N., Liu,B., Liu,Y., Li,W., Li,W., Li,Y.,
            Luo,J., Niu,Y., Qi,Q., Qi,X., Song,S., Sun,M., Sun,W., Sun,Y.,
            Tao,R., Wang,H., Wang,J., Wang,J., Wang,L., Wang,L., Wang,R.,
            Wang,X., Wang,X., Wang,Y., Wu,D., Wu,Q., Xie,F., Xuan,Z., Xue,Y.,
            Yan,C., Yang,X., Yu,B., Zeng,Y., Zhang,G., Zhang,H., Zhang,H.,
            Zhang,L., Zhang,M., Zhang,X., Zhang,X., Zhang,Y., Zhang,Y.,
            Zhang,Z., Zhu,B., Yu,J. and Yang,H.
```

```
  TITLE     Chromosome 3p genomic sequence
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 165379)
  AUTHORS   Wang,L., Hu,S., Dong,W., Wang,J., Zhang,Y., Zhang,H., Liu,B.,
            Bao,W., Sun,Y., Wu,Q., Wang,H., Yang,X., Cheng,C., Wang,Y., Niu,Y.,
            Qi,X., Li,T., Zhang,H., Liu,N., Wu,D., Yu,B., Fan,H., Liu,Y.,
            Li,G., Li,C., Bao,Q., Bao,J., Wang,X., Song,L., Zhang,L., Guo,D.,
            Huang,F., Zhang,G., Li,J., Bian,X., Zhang,M., Li,L., Feng,X., Yu,J.
            and Yang,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-MAY-2000) Human Genomic Center, Institute of
            Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing
            100101, P.R.China
COMMENT     On May 29, 2000 this sequence version replaced gi:8039668.
            --------------Genome Center
            Center:Beijing Center
            Center code:Beijing
            Website:http://hgc.igtp.ac.cn
            http://www.genomics.org.cn
            Contact:hgc@igtp.ac.cn
            -------------- Project Information
            Center project name:1% project
            Center clone name: RP11-208K18
            -------------- Summary Statistics
            Sequencing vector: pUC18; 100% of reads
            Chemistry: Dye-terminator: ET 55% of reads
            Chemistry: Dye-terminator Big Dye; 45% of reads
            Assembly program: Phrap; version 0.990329
            Consensus quality: 164602 bases at least Q40
            Consensus quality: 165885 bases at least Q30
            Consensus quality: 166968 bases at least Q20
            Insert size: 162389; sum-of-contigs
            Quality coverage:   6.82x in Q20 bases;sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 7 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1874: contig of 1874 bp in length
            *     1875     1974: gap of unknown length
            *     1975     8441: contig of 6467 bp in length
            *     8442     8541: gap of unknown length
            *     8542    20983: contig of 12442 bp in length
            *    20984    21083: gap of unknown length
            *    21084    38290: contig of 17207 bp in length
            *    38291    38390: gap of unknown length
            *    38391    56760: contig of 18370 bp in length
            *    56761    56860: gap of unknown length
            *    56861    93526: contig of 36666 bp in length
            *    93527    93626: gap of unknown length
            *    93627   165379: contig of 71753 bp in length.
FEATURES             Location/Qualifiers
     source          1. .165379
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="3"
                     /map="3p"
                     /clone="RP11-208K18"
     misc_feature    1. .1874
                     /note="assembly_name:Contig3"
     misc_feature    1975. .8441
                     /note="assembly_name:Contig4"
     misc_feature    8542. .20983
                     /note="assembly_name:Contig5"
     misc_feature    21084. .38290
                     /note="assembly_name:Contig6"
     misc_feature    38391. .56760
                     /note="assembly_name:Contig7"
     misc_feature    56861. .93526
                     /note="assembly_name:Contig8"
     misc_feature    93627. .165379
                     /note="assembly_name:Contig9"
BASE COUNT    51415 a  32065 c  32897 g  48394 t    608 others
ORIGIN


  Query Match             1.3%;  Score 37;  DB 2;  Length 165379;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-09;
  Matches   37;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   2462 attgagaacctgtcttcccatttggtgtgctttcctc 2498
           |||||||||||||||||||||||||||||||||||
Db  83850 ATTGAGAACCTGTCTTCCCATTTGGTGTGCTTTCCTC 83886


RESULT  13
AC026219
LOCUS       AC026219              184375 bp    DNA     linear   HTG 29-MAY-2000
DEFINITION  Homo sapiens chromosome 3 clone RP11-815J19 map 3p, WORKING DRAFT
            SEQUENCE, 56 unordered pieces.
ACCESSION   AC026219
VERSION     AC026219.2  GI:8101287
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 184375)
  AUTHORS   Bao,J., Bao,Q., Bao,W., Bian,X., Cao,T., Chen,C., Chen,J., Ding,H.,
            Dong,W., Fan,H., Feng,X., Guan,Q., Gu,X., Guo,D., He,L., Hu,S.,
            Huang,F., Jin,Y., Kang,N., Li,C., Li,C., Li,G., Li,J., Li,L.,
            Li,S., Li,T., Liu,Y., Liu,N., Liu,B., Liu,Y., Li,W., Li,W., Li,Y.,
            Luo,J., Niu,Y., Qi,Q., Qi,X., Song,S., Sun,M., Sun,W., Sun,Y.,
            Tao,R., Wang,H., Wang,J., Wang,J., Wang,L., Wang,L., Wang,R.,
            Wang,X., Wang,X., Wang,Y., Wu,D., Wu,Q., Xie,F., Xuan,Z., Xue,Y.,
            Yan,C., Yang,X., Yu,B., Zeng,Y., Zhang,G., Zhang,H., Zhang,H.,
            Zhang,L., Zhang,M., Zhang,X., Zhang,X., Zhang,Y., Zhang,Y.,
            Zhang,Z., Zhu,B., Yu,J. and Yang,H.
  TITLE     Chromosome 3p genomic sequence
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 184375)
  AUTHORS   Kang,N., Hu,S., Dong,W., Wang,J., Zhang,Y., Zhang,H., Liu,B.,
            Bao,W., Sun,Y., Wu,Q., Wang,H., Yang,X., Cheng,C., Wang,Y., Niu,Y.,
            Qi,X., Li,T., Zhang,H., Liu,N., Wu,D., Yu,B., Fan,H., Liu,Y.,
            Li,G., Li,C., Bao,Q., Bao,J., Wang,X., Song,L., Zhang,L., Guo,D.,
            Huang,F., Zhang,G., Li,J., Bian,X., Zhang,M., Li,L., Feng,X., Yu,J.
            and Yang,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-MAR-2000) Human Genomic Center, Institute of
            Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing
            100101, P.R.China
COMMENT     On May 29, 2000 this sequence version replaced gi:7272036.
            --------------Genome Center
            Center:Beijing Center
            Center code:Beijing
            Website:http://hgc.igtp.ac.cn
            http://www.genomics.org.cn
            Contact:hgc@igtp.ac.cn
            -------------- Project Information
            Center project name:1% project
            Center clone name: RP11-815J19
            -------------- Summary Statistics
            Sequencing vector: pUC18; 100% of reads
            Chemistry: Dye-terminator: ET 55% of reads
            Chemistry: Dye-terminator Big Dye; 45% of reads
            Assembly program: Phrap; version 0.990329
            Consensus quality: 146985 bases at least Q40
            Consensus quality: 167323 bases at least Q30
            Consensus quality: 182494 bases at least Q20
            Insert size: 122192; sum-of-contigs
            Quality coverage:   3.83x in Q20 bases;sum-of-contigs
            --------------.
```

```
* NOTE: This is a 'working draft' sequence. It currently
* consists of 56 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*        1     1389: contig of 1389 bp in length
*     1390     1489: gap of unknown length
*     1490     3602: contig of 2113 bp in length
*     3603     3702: gap of unknown length
*     3703     6147: contig of 2445 bp in length
*     6148     6247: gap of unknown length
*     6248     7841: contig of 1594 bp in length
*     7842     7941: gap of unknown length
*     7942     9569: contig of 1628 bp in length
*     9570     9669: gap of unknown length
*     9670    11970: contig of 2301 bp in length
*    11971    12070: gap of unknown length
*    12071    13209: contig of 1139 bp in length
*    13210    13309: gap of unknown length
*    13310    14906: contig of 1597 bp in length
*    14907    15006: gap of unknown length
*    15007    16915: contig of 1909 bp in length
*    16916    17015: gap of unknown length
*    17016    18655: contig of 1640 bp in length
*    18656    18755: gap of unknown length
*    18756    20974: contig of 2219 bp in length
*    20975    21074: gap of unknown length
*    21075    22409: contig of 1335 bp in length
*    22410    22509: gap of unknown length
*    22510    24337: contig of 1828 bp in length
*    24338    24437: gap of unknown length
*    24438    25826: contig of 1389 bp in length
*    25827    25926: gap of unknown length
*    25927    27769: contig of 1843 bp in length
*    27770    27869: gap of unknown length
*    27870    29441: contig of 1572 bp in length
*    29442    29541: gap of unknown length
*    29542    32494: contig of 2953 bp in length
*    32495    32594: gap of unknown length
*    32595    34633: contig of 2039 bp in length
*    34634    34733: gap of unknown length
*    34734    37159: contig of 2426 bp in length
*    37160    37259: gap of unknown length
*    37260    39230: contig of 1971 bp in length
*    39231    39330: gap of unknown length
*    39331    40642: contig of 1312 bp in length
*    40643    40742: gap of unknown length
*    40743    42739: contig of 1997 bp in length
*    42740    42839: gap of unknown length
*    42840    45245: contig of 2406 bp in length
*    45246    45345: gap of unknown length
*    45346    47083: contig of 1738 bp in length
*    47084    47183: gap of unknown length
*    47184    49945: contig of 2762 bp in length
*    49946    50045: gap of unknown length
*    50046    51544: contig of 1499 bp in length
*    51545    51644: gap of unknown length
*    51645    53891: contig of 2247 bp in length
*    53892    53991: gap of unknown length
*    53992    57221: contig of 3230 bp in length
*    57222    57321: gap of unknown length
*    57322    60016: contig of 2695 bp in length
*    60017    60116: gap of unknown length
*    60117    62563: contig of 2447 bp in length
*    62564    62663: gap of unknown length
*    62664    64926: contig of 2263 bp in length
*    64927    65026: gap of unknown length
*    65027    68373: contig of 3347 bp in length
*    68374    68473: gap of unknown length
*    68474    71700: contig of 3227 bp in length
*    71701    71800: gap of unknown length
*    71801    74355: contig of 2555 bp in length
*    74356    74455: gap of unknown length
*    74456    77266: contig of 2811 bp in length
*    77267    77366: gap of unknown length
*    77367    80106: contig of 2740 bp in length
*    80107    80206: gap of unknown length
*    80207    82763: contig of 2557 bp in length
*    82764    82863: gap of unknown length
*    82864    85878: contig of 3015 bp in length
*    85879    85978: gap of unknown length
*    85979    88748: contig of 2770 bp in length
*    88749    88848: gap of unknown length
*    88849    93166: contig of 4318 bp in length
*    93167    93266: gap of unknown length
*    93267    96634: contig of 3368 bp in length
*    96635    96734: gap of unknown length
*    96735   100449: contig of 3715 bp in length
*   100450   100549: gap of unknown length
*   100550   103083: contig of 2534 bp in length
*   103084   103183: gap of unknown length
*   103184   107029: contig of 3846 bp in length
*   107030   107129: gap of unknown length
*   107130   110324: contig of 3195 bp in length
*   110325   110424: gap of unknown length
*   110425   114467: contig of 4043 bp in length
*   114468   114567: gap of unknown length
*   114568   119345: contig of 4778 bp in length
*   119346   119445: gap of unknown length
*   119446   122439: contig of 2994 bp in length
*   122440   122539: gap of unknown length
*   122540   127598: contig of 5059 bp in length
*   127599   127698: gap of unknown length
*   127699   133792: contig of 6094 bp in length
*   133793   133892: gap of unknown length
*   133893   140225: contig of 6333 bp in length
*   140226   140325: gap of unknown length
*   140326   144336: contig of 4011 bp in length
*   144337   144436: gap of unknown length
*   144437   151171: contig of 6735 bp in length
*   151172   151271: gap of unknown length
*   151272   159310: contig of 8039 bp in length
*   159311   159410: gap of unknown length
*   159411   168312: contig of 8902 bp in length
*   168313   168412: gap of unknown length
*   168413   184375: contig of 15963 bp in length.
FEATURES             Location/Qualifiers
     source          1. .184375
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="3"
                     /map="3p"
                     /clone="RP11-815J19"
     misc_feature    1. .1389
                     /note="assembly_name:Contig35"
     misc_feature    1490. .3602
                     /note="assembly_name:Contig37"
     misc_feature    3703. .6147
                     /note="assembly_name:Contig42"
     misc_feature    6248. .7841
                     /note="assembly_name:Contig45"
     misc_feature    7942. .9569
                     /note="assembly_name:Contig48"
     misc_feature    9670. .11970
                     /note="assembly_name:Contig50"
     misc_feature    12071. .13209
                     /note="assembly_name:Contig51"
     misc_feature    13310. .14906
                     /note="assembly_name:Contig52"
     misc_feature    15007. .16915
                     /note="assembly_name:Contig53"
     misc_feature    17016. .18655
```

```
  Query Match             1.3%;  Score 37;  DB 2;  Length 184375;
  Best Local Similarity   100.0%;  Pred. No. 2.3e-09;
  Matches   37;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    2462 attgagaacctgtcttcccatttggtgtgctttcctc 2498
           |||||||||||||||||||||||||||||||||||||
Db  158390 ATTGAGAACCTGTCTTCCCATTTGGTGTGCTTTCCTC 158426


RESULT  14
AC027579
LOCUS       AC027579              148270 bp    DNA     linear   HTG 23-JUN-2000
DEFINITION  Homo sapiens chromosome 17 clone RP11-177D1 map 17, WORKING DRAFT
            SEQUENCE, 13 unordered pieces.
ACCESSION   AC027579
VERSION     AC027579.2  GI:8671976
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 148270)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 17, clone RP11-177D1
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 148270)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
            Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
            Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J.,
            Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
            McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
            Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
            Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N.,
            Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
            Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
            Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
            Young,G., Zainoun,J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-MAR-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Jun 23, 2000 this sequence version replaced gi:7342324.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L9001
                Center clone name: 177_D_1
            -------------- Summary Statistics
                Sequencing vector: M13; M77815; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 141141 bases at least Q40
                Consensus quality: 144792 bases at least Q30
                Consensus quality: 146242 bases at least Q20
                Insert size: 151000; agarose-fp
                Insert size: 147070; sum-of-contigs
                Quality coverage: 4.7 in Q20 bases; agarose-fp
                Quality coverage: 4.8 in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 13 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1      131: contig of 131 bp in length
            *      132 231: gap of      100 bp
            *      232     4256: contig of 4025 bp in length
            *     4257 4356: gap of      100 bp
            *     4357     6715: contig of 2359 bp in length
            *     6716 6815: gap of      100 bp
            *     6816    12854: contig of 6039 bp in length
            *    12855 12954: gap of      100 bp
            *    12955    18900: contig of 5946 bp in length
            *    18901 19000: gap of      100 bp
            *    19001    27560: contig of 8560 bp in length
            *    27561 27660: gap of      100 bp
            *    27661    41935: contig of 14275 bp in length
            *    41936 42035: gap of      100 bp
            *    42036    58175: contig of 16140 bp in length
            *    58176 58275: gap of      100 bp
            *    58276    70930: contig of 12655 bp in length
            *    70931 71030: gap of      100 bp
            *    71031    87632: contig of 16602 bp in length
            *    87633 87732: gap of      100 bp
            *    87733   104759: contig of 17027 bp in length
            *   104760 104859: gap of      100 bp
            *   104860   123863: contig of 19004 bp in length
            *   123864 123963: gap of      100 bp
            *   123964   148270: contig of 24307 bp in length.
FEATURES             Location/Qualifiers
     source          1. .148270
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="17"
                     /map="17"
                     /clone="RP11-177D1"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .131
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"
     misc_feature    232. .4256
                     /note="assembly_fragment"
     misc_feature    4357. .6715
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:right"
     misc_feature    6816. .12854
                     /note="assembly_fragment"
     misc_feature    12955. .18900
                     /note="assembly_fragment"
     misc_feature    19001. .27560
                     /note="assembly_fragment"
     misc_feature    27661. .41935
                     /note="assembly_fragment"
     misc_feature    42036. .58175
                     /note="assembly_fragment"
     misc_feature    58276. .70930
                     /note="assembly_fragment"
     misc_feature    71031. .87632
                     /note="assembly_fragment"
     misc_feature    87733. .104759
                     /note="assembly_fragment"
     misc_feature    104860. .123863
                     /note="assembly_fragment"
     misc_feature    123964. .148270
                     /note="assembly_fragment"
BASE COUNT  40333 a  32099 c  32432 g  42206 t   1200 others
```

```
ORIGIN


  Query Match             1.3%;  Score 36;  DB 2;  Length 148270;
  Best Local Similarity   100.0%;  Pred. No. 9.1e-09;
  Matches   36;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   2462 attgagaacctgtcttcccatttggtgtgctttcct 2497
          ||||||||||||||||||||||||||||||||||||
Db  76167 ATTGAGAACCTGTCTTCCCATTTGGTGTGCTTTCCT 76202


RESULT  15
AC022727
LOCUS       AC022727               174405 bp    DNA     linear   HTG 26-MAY-2000
DEFINITION  Homo sapiens chromosome 18 clone RP11-308J14 map 18, WORKING DRAFT
            SEQUENCE, 11 unordered pieces.
ACCESSION   AC022727
VERSION     AC022727.4  GI:8072614
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 174405)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 18, clone RP11-308J14
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 174405)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
            DeArellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
            Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
            Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
            Landers,T., Lehoczky,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
            Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
            McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J.,
            Norman,C.H., O'Connor,T., O'Donnell,P., Olivar,T.M., Peterson,K.,
            Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
            Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
            Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-FEB-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On May 25, 2000 this sequence version replaced gi:7658384.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L5488
                Center clone name: 308_J_14
            -------------- Summary Statistics
                Sequencing vector: M13; M77815; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 169089 bases at least Q40
                Consensus quality: 172164 bases at least Q30
                Consensus quality: 173032 bases at least Q20
                Insert size: 177000; agarose-fp
                Insert size: 173405; sum-of-contigs
                Quality coverage: 4.8 in Q20 bases; agarose-fp
                Quality coverage: 4.9 in Q20 bases; sum-of-contigs
            --------------
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 11 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     6450: contig of 6450 bp in length
            *     6451 6550: gap of      100 bp
            *     6551    12911: contig of 6361 bp in length
            *    12912 13011: gap of      100 bp
            *    13012    22057: contig of 9046 bp in length
            *    22058 22157: gap of      100 bp
            *    22158    32841: contig of 10684 bp in length
            *    32842 32941: gap of      100 bp
            *    32942    48367: contig of 15426 bp in length
            *    48368 48467: gap of      100 bp
            *    48468    63919: contig of 15452 bp in length
            *    63920 64019: gap of      100 bp
            *    64020    78259: contig of 14240 bp in length
            *    78260 78359: gap of      100 bp
            *    78360    98415: contig of 20056 bp in length
            *    98416 98515: gap of      100 bp
            *    98516   121876: contig of 23361 bp in length
            *   121877 121976: gap of      100 bp
            *   121977   144489: contig of 22513 bp in length
            *   144490 144589: gap of      100 bp
            *   144590   174405: contig of 29816 bp in length.
FEATURES             Location/Qualifiers
     source          1. .174405
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="18"
                     /map="18"
                     /clone="RP11-308J14"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .6450
                     /note="assembly_fragment"
     misc_feature    6551. .12911
                     /note="assembly_fragment"
     misc_feature    13012. .22057
                     /note="assembly_fragment"
     misc_feature    22158. .32841
                     /note="assembly_fragment"
     misc_feature    32942. .48367
                     /note="assembly_fragment"
     misc_feature    48468. .63919
                     /note="assembly_fragment"
     misc_feature    64020. .78259
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:right"
     misc_feature    78360. .98415
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:left"
     misc_feature    98516. .121876
                     /note="assembly_fragment"
     misc_feature    121977. .144489
                     /note="assembly_fragment"
     misc_feature    144590. .174405
                     /note="assembly_fragment"
BASE COUNT    46895 a  38223 c  37328 g  50956 t   1003 others
ORIGIN


  Query Match             1.3%;  Score 36;  DB 2;  Length 174405;
  Best Local Similarity   100.0%;  Pred. No. 9e-09;
  Matches   36;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   2462 attgagaacctgtcttcccatttggtgtgctttcct 2497
          ||||||||||||||||||||||||||||||||||||
```

Db 127497 ATTGAGAACCTGTCTTCCCATTTGGTGTGCTTTCCT 127532

Search completed: May 17, 2002, 11:55:34
Job time: 75409 sec

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: May 16, 2002, 14:43:43 ; Search time 539.13 Seconds
(without alignments)
8827.717 Million cell updates/sec

Title: US-09-898-556A-3
Perfect score: 2772
Sequence: 1 caggcgcgttaagctggttg.....tttctacccatcctcaccct 2772

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1736436 seqs, 858457221 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_032802:*
1: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1980.DAT:*
2: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1981.DAT:*
3: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1982.DAT:*
4: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1983.DAT:*
5: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1984.DAT:*
6: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1985.DAT:*
7: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1986.DAT:*
8: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1987.DAT:*
9: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1988.DAT:*
10: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1989.DAT:*
11: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1990.DAT:*
12: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1991.DAT:*
13: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1992.DAT:*
14: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1993.DAT:*
15: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1994.DAT:*
16: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1995.DAT:*
17: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1996.DAT:*
18: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1997.DAT:*
19: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1998.DAT:*
20: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1999.DAT:*
21: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2000.DAT:*
22: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:*
23: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001B.DAT:*
24: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1577	56.9	2614	23	AAS74823	DNA encoding novel
2	571	20.6	678	22	AAF44909	Human breast cance
3	525	18.9	653	23	AAS68256	DNA encoding novel
4	231	8.3	983	22	AAS26286	Human cDNA encodin
5	99	3.6	394	22	AAI83890	Human polynucleoti
6	69	2.5	69	22	AAS13031	DNA encoding zinc
7	69	2.5	69	22	AAS13088	DNA encoding zinc
8	49	1.8	821	23	AAS68257	DNA encoding novel
9	29	1.0	800	22	AAH33832	Human colon cancer
10	29	1.0	2349	23	AAS90429	DNA encoding novel
11	29	1.0	3582	22	AAI58664	Human polynucleoti
c 12	29	1.0	3582	22	AAI60450	Human polynucleoti
13	28	1.0	159	22	AAL00216	Human reproductive
14	27	1.0	1183	22	AAS25949	Human cDNA encodin
15	27	1.0	2717	22	AAS26635	Human genomic DNA
16	27	1.0	3099	20	AAV64361	Human stem cell zi
c 17	25	0.9	51	23	ABL00474	Human silent nonco
c 18	25	0.9	51	23	ABL00878	Human amino acid c
19	25	0.9	763	22	AAK92290	Human cDNA 5'-end
20	25	0.9	763	22	AAK93811	Human cDNA clone r
21	25	0.9	814	22	AAI95339	Human neuroblastom
22	25	0.9	871	22	ABA08823	Human PRO1847 homo
23	25	0.9	2200	22	AAK94575	Human full-length
24	23	0.8	688	22	AAK91560	Human cDNA 5'-end
25	23	0.8	688	22	AAK93129	Human cDNA clone r
c 26	23	0.8	1512	22	ABA50438	Human breast cell
c 27	23	0.8	1512	22	ABA68388	Human foetal liver
c 28	23	0.8	1512	22	ABA35382	Probe #13848 for g
c 29	23	0.8	1512	22	AAK16761	Human brain expres
c 30	23	0.8	1512	22	AAK42537	Human bone marrow
c 31	23	0.8	1512	22	AAI23284	Probe #13217 for g
c 32	23	0.8	1512	22	AAI48604	Probe #17290 used
c 33	23	0.8	1512	22	AAI08925	Probe #8916 used t
34	23	0.8	1623	22	AAK94635	Human full-length
c 35	23	0.8	1965	22	ABA45305	Human breast cell
c 36	23	0.8	1965	22	ABA55794	Human foetal liver
c 37	23	0.8	1965	22	ABA25474	Probe #3940 for ge
c 38	23	0.8	1965	22	AAK04015	Human brain expres
c 39	23	0.8	1965	22	AAK29501	Human bone marrow
c 40	23	0.8	1965	22	AAI14072	Probe #4005 for ge
c 41	23	0.8	1965	22	AAI35452	Probe #4138 used t
c 42	23	0.8	1965	22	AAI03925	Probe #3916 used t
43	23	0.8	2634	23	AAS81041	DNA encoding novel
44	22	0.8	439	22	AAF67397	Novel human polynu
45	22	0.8	484	21	AAC79234	Human lung tumour-

ALIGNMENTS

RESULT 1
AAS74823
ID AAS74823 standard; cDNA; 2614 BP.
XX
AC AAS74823;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #10627.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG10636.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 1; SEQ ID No 10627; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2614 BP; 745 A; 611 C; 660 G; 598 T; 0 other;

Query Match 56.9%; Score 1577; DB 23; Length 2614;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1817; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

```
Qy    427 cagaatcgaagccagaaattcaacttagtccctcctgccctctgattttctccagtcagc 486
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    229 cagaatcgaagccagaaattcaacttagtccctcctgccctctgattttctccagtcagc 288

Qy    487 aagctctcagccaacatgtgtggctgagtcatctctctcagctgttttcaagtttatggg 546
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    289 aagctctcagccaacatgtgtggctgagtcatctctctcagctgttttcaagtttatggg 348

Qy    547 cagg-aaatcctctccacctgggaaaacactatccagaagatcagaaacaacagcaggat 605
          |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    349 caggaaaatcctctccacctgggaaaacactatccagaagatcagaaacaacagcaggat 408

Qy    606 ccattctgctttagtggcaaagcagaatggattcaagagggagaagactccagactcctg 665
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    409 ccattctgctttagtggcaaagcagaatggattcaagagggagaagactccagactcctg 468

Qy    666 tttgggagagtaagcaaaaatggcacttcaaaggcactttccagcccacctgaagaacaa 725
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    469 tttgggagagtaagcaaaaatggcacttcaaaggcactttccagcccacctgaagaacaa 528

Qy    726 cagccagcacagtccaaggaagacaacacagtggtggatatagggtccagccctgaacgg 785
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    529 cagccagcacagtccaaggaagacaacacagtggtggatatagggtccagccctgaacgg 588

Qy    786 agggcagatctagaggaaacagacaaagtattgcatggtttagaagtctcaggatttgga 845
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    589 agggcagatctagaggaaacagacaaagtattgcatggtttagaagtctcaggatttgga 648

Qy    846 gaaatcaaatatgaagagtttgggccaggctttatcaaggagtcaaacctccttagcctc 905
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    649 gaaatcaaatatgaagagtttgggccaggctttatcaaggagtcaaacctccttagcctc 708

Qy    906 cagaagacacaaactggggagacaccttacatgtacactgagtggggagacagctttggc 965
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    709 cagaagacacaaactggggagacaccttacatgtacactgagtggggagacagctttggc 768

Qy    966 agtatgtcagtcctcatcaaaaacccaaggacacactctgggggaaagccttatgtgtgc 1025
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db    769 agtatgtcagtcctcatcaaaaacccaaggacacactctgggggaaagccttatgtgtgc 828

Qy   1026 agggaatgtgggcgaggctttacgtggaagtcaaacctgatcacacatcagaggacacac 1085
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    829 agggaatgtgggcgaggctttacgtggaagtcaaacctgatcacacatcagaggacacac 888

Qy   1086 tcaggggagaaaccttatgtgtgcaaggattgtggacgaggctttacttggaagtcgaac 1145
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    889 tcaggggagaaaccttatgtgtgcaaggattgtggacgaggctttacttggaagtcgaac 948

Qy   1146 ctctttacacatcagcggacacactcagggctcaagccttatgtgtgcaaggaatgtggg 1205
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    949 ctctttacacatcagcggacacactcagggctcaagccttatgtgtgcaaggaatgtggg 1008

Qy   1206 cagagctttagcctgaagtcaaacctcattacccaccagagggcgcacactggggagaag 1265
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1009 cagagctttagcctgaagtcaaacctcattacccaccagagggcgcacactggggagaag 1068

Qy   1266 ccttatgtttgcagggaatgtgggcgtggctttcgccagcattcacacctggtcagacac 1325
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1069 ccttatgtttgcagggaatgtgggcgtggctttcgccagcattcacacctggtcagacac 1128

Qy   1326 aagaggacacattcaggagagaagccttacatttgcagggagtgtgagcaaggctttagc 1385
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1129 aagaggacacattcaggagagaagccttacatttgcagggagtgtgagcaaggctttagc 1188

Qy   1386 cagaagtcacacctcatcagacacttaaggacacacacaggagagaagccttatgtatgc 1445
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1189 cagaagtcacacctcatcagacacttaaggacacacacaggagagaagccttatgtatgc 1248

Qy   1446 acagaatgtgggcgtcactttagctggaaatcaaacctcaaaacacaccagaggacacac 1505
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1249 acagaatgtgggcgtcactttagctggaaatcaaacctcaaaacacaccagaggacacac 1308

Qy   1506 tcaggggttaaaccttatgtctgcctggagtgcgggcagtgctttagcctgaagtcaaac 1565
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1309 tcaggggttaaaccttatgtctgcctggagtgcgggcagtgctttagcctgaagtcaaac 1368

Qy   1566 cttaacaaacacc-agaggtcacacacgggggagaagccatttgtatgtacggagtgtgg 1624
          |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1369 cttaacaaacaccaagaggtcacacacgggggagaagccatttgtatgtacggagtgtgg 1428

Qy   1625 gcgaggctttacccggaaatcaaccctgatcacgcaccagaggacacactcaggggagaa 1684
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1429 gcgaggctttacccggaaatcaaccctgatcacgcaccagaggacacactcaggggagaa 1488

Qy   1685 gccatttgtatgtgctgagtgtggacgaggctttaatgataagtccaccctcatttcaca 1744
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1489 gccatttgtatgtgctgagtgtggacgaggctttaatgataagtccaccctcatttcaca 1548

Qy   1745 ccagaggacacattcaggggaaaagccttttatgtgcagggagtgtggcagaaggtttcg 1804
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1549 ccagaggacacattcaggggaaaagccttttatgtgcagggagtgtggcagaaggtttcg 1608

Qy   1805 gcagaagcctaacctgtttaggcacaagagggcacactcaggtgcctttgtgtgcaggga 1864
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1609 gcagaagcctaacctgtttaggcacaagagggcacactcaggtgcctttgtgtgcaggga 1668

Qy   1865 gtgtgggcaaggcttttgtgctaagttaactctcattaaacaccagagagcacacgcagg 1924
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1669 gtgtgggcaaggcttttgtgctaagttaactctcattaaacaccagagagcacacgcagg 1728

Qy   1925 ggggaagcctcatgtgtgcagggagtgtgggcaaggctttagccggcagtcacacctcat 1984
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1729 ggggaagcctcatgtgtgcagggagtgtgggcaaggctttagccggcagtcacacctcat 1788

Qy   1985 tagacaccagaggacacattcaggagagaagccttatatttgcagaaagtgtggacgggg 2044
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1789 tagacaccagaggacacattcaggagagaagccttatatttgcagaaagtgtggacgggg 1848

Qy   2045 ctttagtcggaagtccaaccttatcagacatcagaggacacactcaggatagaaacttta 2104
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1849 ctttagtcggaagtccaaccttatcagacatcagaggacacactcaggatagaaacttta 1908
```

```
Qy   2105 tgtgtatagggaatgtggtacagcctttagccaggagtcatacttcatcagacaccagag 2164
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1909 tgtgtatagggaatgtggtacagcctttagccaggagtcatacttcatcagacaccagag 1968

Qy   2165 gacacacacagtgctgtggctttttcagccattgctagataccaaagtggagacattctg 2224
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1969 gacacacacagtgctgtggctttttcagccattgctagataccaaagtggagacattctg 2028

Qy   2225 tgtgtgattatgcatgaga 2243
          ||||||||||||||||||
Db   2029 tgtgtgattatgcatgaga 2047


RESULT   2
AAF44909
ID   AAF44909 standard; cDNA; 678 BP.
XX
AC   AAF44909;
XX
DT   28-MAR-2001  (first entry)
XX
DE   Human breast cancer related protein coding sequence SEQ ID NO: 65.
XX
KW   Human; breast cancer; diagnosis; therapy; vaccine; ss.
XX
OS   Homo sapiens.
XX
PN   WO200078960-A2.
XX
PD   28-DEC-2000.
XX
PF   23-JUN-2000; 2000WO-US17536.
XX
PR   23-JUN-1999;   99US-0140903.
PR   12-OCT-1999;   99US-0158980.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Yuqiu J,  Mitcham JL;
XX
DR   WPI; 2001-041426/05.
XX
PT   New polynucleotides encoding breast tumor specific proteins, useful for
PT   prevention, treatment and diagnosis of breast cancer -
XX
PS   Claim 25; Page 141; 165pp; English.
XX
CC   The present invention provides the coding sequences for a number of
CC   breast cancer related proteins. These can be used in vaccinations
CC   against, diagnosis of and treatment of cancer, particularly breast
CC   cancer.
XX
SQ   Sequence 678 BP; 165 A; 167 C; 160 G; 182 T; 4 other;


  Query Match             20.6%;  Score 571;  DB 22;  Length 678;
  Best Local Similarity   100.0%;  Pred. No. 7.2e-271;
  Matches  571;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   2078 gaggacacactcaggatagaaactttatgtgtatagggaatgtggtacagcctttagcca 2137
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      9 gaggacacactcaggatagaaactttatgtgtatagggaatgtggtacagcctttagcca 68

Qy   2138 ggagtcatacttcatcagacaccagaggacacacacagtgctgtggctttttcagccatt 2197
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     69 ggagtcatacttcatcagacaccagaggacacacacagtgctgtggctttttcagccatt 128

Qy   2198 gctagataccaaagtggagacattctgtgtgtgattatgcatgagactgtactggtaaga 2257
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    129 gctagataccaaagtggagacattctgtgtgtgattatgcatgagactgtactggtaaga 188

Qy   2258 cttgtatctccatccacctgaaggagaattgctggctcattttcaggagccctgcccttc 2317
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    189 cttgtatctccatccacctgaaggagaattgctggctcattttcaggagccctgcccttc 248

Qy   2318 ctcactgtggatggtgggttgtggaaacccggtcaggtaatgatagtggcaggaggcagt 2377
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    249 ctcactgtggatggtgggttgtggaaacccggtcaggtaatgatagtggcaggaggcagt 308

Qy   2378 caaatgcccaggcagataggggtgggtacctggtgaaacccaaccttaaagctgaagaca 2437
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    309 caaatgcccaggcagataggggtgggtacctggtgaaacccaaccttaaagctgaagaca 368

Qy   2438 gtcccggctaaatcctcatactgaattgagaacctgtcttcccatttggtgtgctttcct 2497
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    369 gtcccggctaaatcctcatactgaattgagaacctgtcttcccatttggtgtgctttcct 428

Qy   2498 ccgattgatcccaacccttcacctattttacgtatacctgccctttcctaattggttttt 2557
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    429 ccgattgatcccaacccttcacctattttacgtatacctgccctttcctaattggttttt 488

Qy   2558 acactgctgtgcccacctttgagtggtgcctttgcatacttacaaatcagtcaacgtgt 2617
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    489 acactgctgtgcccacctttgagtggtgcctttgcatacttacaaatcagtcaacgtgt 548

Qy   2618 attcccctattctgagcccataaaagaccca 2648
          |||||||||||||||||||||||||||||
Db    549 attcccctattctgagcccataaaagaccca 579


RESULT   3
AAS68256
ID   AAS68256 standard; cDNA; 653 BP.
XX
AC   AAS68256;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #4060.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG04069.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity  -
XX
PS   Claim 1; SEQ ID No 4060; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
```

```
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human
CC   diagnostic coding sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 653 BP; 192 A; 151 C; 176 G; 134 T; 0 other;

  Query Match             18.9%;  Score 525;  DB 23;  Length 653;
  Best Local Similarity   99.8%;  Pred. No. 3.3e-248;
  Matches  575;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy    1004 tgggggaaagccttatgtgtgcagggaatgtgggcgaggctttacgtggaagtcaaacct 1063
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      52 tgggggaaagccttatgtgtgcagggaatgtgggcgaggctttacgtggaagtcaaacct 111

Qy    1064 gatcacacatcagaggacacactcaggggagaaaccttatgtgtgcaaggattgtggacg 1123
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     112 gatcacacatcagaggacacactcaggggagaaaccttatgtgtgcaaggattgtggacg 171

Qy    1124 aggctttacttggaagtcgaacctctttacacatcagcggacacactcagggctcaagcc 1183
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     172 aggctttacttggaagtcgaacctctttacacatcagcggacacactcagggctcaagct 231

Qy    1184 ttatgtgtgcaaggaatgtgggcagagctttagcctgaagtcaaacctcattacccacca 1243
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     232 ttatgtgtgcaaggaatgtgggcagagctttagcctgaagtcaaacctcattacccacca 291

Qy    1244 gagggcgcacactggggagaagccttatgtttgcagggaatgtgggcgtggctttcgcca 1303
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     292 gagggcgcacactggggagaagccttatgtttgcagggaatgtgggcgtggctttcgcca 351

Qy    1304 gcattcacacctggtcagacacaagaggacacattcaggagagaagccttacatttgcag 1363
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     352 gcattcacacctggtcagacacaagaggacacattcaggagagaagccttacatttgcag 411

Qy    1364 ggagtgtgagcaaggctttagccagaagtcacacctcatcagacacttaaggacacacac 1423
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     412 ggagtgtgagcaaggctttagccagaagtcacacctcatcagacacttaaggacacacac 471

Qy    1424 aggagagaagccttatgtatgcacagaatgtgggcgtcactttagctggaaatcaaacct 1483
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     472 aggagagaagccttatgtatgcacagaatgtgggcgtcactttagctggaaatcaaacct 531

Qy    1484 caaaacacaccagaggacacactcaggggttaaaccttatgtctgcctggagtgcgggca 1543
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     532 caaaacacaccagaggacacactcaggggttaaaccttatgtctgcctggagtgcgggca 591

Qy    1544 gtgctttagcctgaagtcaaaccttaacaaacacca 1579
           |||||||||||||||||||||||||||||||||||
Db     592 gtgctttagcctgaagtcaaaccttaacaaacacca 627


RESULT   4
AAS26286
ID   AAS26286 standard; cDNA; 983 BP.
XX
AC   AAS26286;
XX
DT   07-NOV-2001  (first entry)
XX
DE   Human cDNA encoding a novel secreted protein, Seq ID 465.
XX
KW   Human; immunosuppressive; antiarthritic; ss; antirheumatic;
KW   cytostatic; cardiant; vasotropic; cerebroprotective; nootropic;
KW   neuroprotective; antibacterial; virucide; fungicide; opthalmalogical;
KW   vulnerary; secreted protein; rheumatoid arthritis;
KW   hyperproliferative disorder; cardiovascular disorder; cardiac arrest;
KW   cerebrovascular disorder; cerebral ischaemia; angiogenesis;
KW   nervous system disorder; Alzheimer's disease; infection; ocular disorder;
KW   corneal infection; wound healing; epithelial cell proliferation;
KW   skin ageing; food additive; preservative; antiproliferative.
XX
OS   Homo sapiens.
XX
PN   WO200155322-A2.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01341.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
PR   08-SEP-2000; 2000US-0232081.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232397.
PR   14-SEP-2000; 2000US-0232398.
PR   14-SEP-2000; 2000US-0232399.
```

PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-488783/53.
DR P-PSDB; AAU16299.
XX
PT New nucleic acid molecules encoding 461 human secreted proteins for
PT diagnosing, preventing, treating or ameliorating medical conditions and
PT used as food additives or preservatives -
XX
PS Claim 1; SEQ ID No 465; 980pp; English.
XX
CC The invention relates to isolated nucleic acid molecules and their
CC encoded secreted proteins. The nucleic acids and proteins are used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They
CC are also used in diagnosing a pathological condition or susceptibility
CC to a pathological condition. Antibodies to the proteins can also
CC be used in alleviating symptoms associated with the disorders and in
CC diagnostic immunoassays e.g. radioimmunoassays or enzyme linked
CC immunosorbant assays (ELISA). Disorders which are diagnosed or treated
CC include autoimmune diseases e.g. rheumatoid arthritis,
CC hyperproliferative disorders e.g. neoplasms of the breast or liver,
CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
CC e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g.
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi
CC and ocular disorders e.g. corneal infection, and many other
CC disorders listed in the specification. The polypeptides can also
CC be used to aid wound healing and epithelial cell proliferation, to
CC prevent skin aging due to sunburn, to maintain organs before
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The polypeptides can also be used
CC as a food additive or preservative to increase or decrease storage
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
CC minerals, cofactors and other nutritional components. The present
CC sequence encodes a novel secreted protein of the invention.

Query Match 8.3%; Score 231; DB 22; Length 983;
Best Local Similarity 99.6%; Pred. No. 2.4e-103;
Matches 281; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy    287 gcacagggaggtgatgctggagacttataaccatctggtctcactggaaattccatcttc 346
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    570 gcacagggaggtgatgctggagacttataaccatctggtctcactggaaattccatcttc 629

Qy    347 taaaccaaaactcattgctcagctggagcgaggggaagcgccctggagagaggagagaaa 406
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    630 taaaccaaaactcattgctcagctggagcgaggggaagcgccctggagagaggagagaaa 689

Qy    407 atgtccactggacctctgtccagaatcgaagccagaaattcaacttagtccctcctgccc 466
          |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db    690 atgtccactggacctctgtccanaatcgaagccagaaattcaacttagtccctcctgccc 749

Qy    467 tctgattttctccagtcagcaagctctcagccaacatgtgtggctgagtcatctctctca 526
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    750 tctgattttctccagtcagcaagctctcagccaacatgtgtggctgagtcatctctctca 809

Qy    527 gctgttttcaagtttatgggcaggaaatcctctccacctggg 568
          |||||||||||||||||||||||||||||||||||||||||
Db    810 gctgttttcaagtttatgggcaggaaatcctctccacctggg 851
```

RESULT 5

```
AAI83890
ID   AAI83890 standard; cDNA; 394 BP.
XX
AC   AAI83890;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human polynucleotide SEQ ID NO 3950.
XX
KW   Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW   vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW   tissue growth factor; immunomodulatory; cancer; leukaemia;
KW   nervous system disorders; arthritis; inflammation; ss.
XX
OS   Homo sapiens.
XX
PN   WO200164835-A2.
XX
PD   07-SEP-2001.
XX
PF   26-FEB-2001; 2001WO-US04927.
XX
PR   28-FEB-2000; 2000US-0515126.
PR   18-MAY-2000; 2000US-0577409.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Drmanac RT;
XX
DR   WPI; 2001-514838/56.
DR   P-PSDB; AAO03959.
XX
PT   Isolated nucleic acids and polypeptides, useful for preventing
PT   diagnosing and treating e.g. leukaemia, inflammation and immune
PT   disorders -
XX
PS   Claim 1; SEQ ID NO 3950; 1399pp + Sequence Listing; English.
XX
CC   The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC   the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
CC   cytokine, cell proliferation or cell differentiation or which may induce
CC   production of other cytokines in other cell populations. The
CC   polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC   peptide therapy. The polypeptides have various cytokine-like activities,
CC   e.g. stem cell growth factor activity, haematopoiesis regulating
CC   activity, tissue growth factor activity, immunomodulatory activity and
CC   activin/inhibin activity and may be useful in the diagnosis and/or
CC   treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC   inflammation.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 394 BP; 148 A; 75 C; 93 G; 78 T; 0 other;


  Query Match             3.6%;  Score 99;  DB 22;  Length 394;
  Best Local Similarity   99.3%;  Pred. No. 2.6e-38;
  Matches  149;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy   2619 ttcccctattctgagcccataaaagacccagactcagctgcagtgaggagagaaatcacc 2678
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ttcccctattctgagcccataaaagacccagactcagctgcagtgaggagagaaatcacc 60

Qy   2679 ctgctgtggaggttggggaccactccctgcatcccctctccactgagagctgttcttttg 2738
           ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ctgctgtgggggttggggaccactccctgcatcccctctccactgagagctgttcttttg 120

Qy   2739 ctcaataaaattcttttctacccatcctca 2768
           |||||||||||||||||||||||||||||
Db    121 ctcaataaaattcttttctacccatcctca 150
```

```
RESULT   6
AAS13031
ID   AAS13031 standard; DNA; 69 BP.
XX
AC   AAS13031;
XX
DT   17-DEC-2001  (first entry)
XX
DE   DNA encoding zinc finger domain TG-ZFD-012.
XX
KW   Zinc finger domain; cancer; human; ds; TG-ZFD-012.
XX
OS   Homo sapiens.
XX
PN   WO200160970-A2.
XX
PD   23-AUG-2001.
XX
PF   17-FEB-2001; 2001WO-KR00244.
XX
PR   18-FEB-2000; 2000KR-0007730.
XX
PA   (TOOL-) TOOLGEN INC.
XX
PI   Kim J,  Kwon Y,  Kim H,  Ryu EH,  Hwang MS;
XX
DR   WPI; 2001-557644/62.
DR   P-PSDB; AAU08514.
XX
PT   Identifying a zinc finger domain for e.g. designing new polypeptides
PT   that bind to a specific site on a DNA, comprises expressing hybrid
PT   nucleic acids with a test zinc finger domain in cells -
XX
PS   Example 22; Page 55; 147pp; English.
XX
CC   The invention relates to a mèthod of identifying a zinc finger domain
CC   that recognises a target site on a DNA. The method comprises expressing
CC   hybrid nucleic acids with a test zinc finger domain in cells containing a
CC   reporter construct, where the reporter gene is expressed above a given
CC   level when a transcription factor recognises a recruitment and a target
CC   site of a promoter, and not only the recruitment site of the promoter.
CC   The method is used to: (a) identify a zinc finger domain that recognises
CC   a target site on a DNA; (b) determine whether a test zinc finger domain
CC   recognises a target site on a promoter; (c) generate a nucleic acid that
CC   encodes a chimeric zinc finger protein; and (d) identify DNA sequences
CC   recognised by zinc finger domains. The method can be used to design novel
CC   polypeptides that bind to a specific site on a DNA. The method can
CC   facilitate the customised generation of new polypeptides that can
CC   regulate the expression a selected target e.g. a gene required by a
CC   pathogen can be repressed, a gene required for cancerous growth can be
CC   repressed, or a gene poorly expressed or encoding a mutated protein can
CC   be activated and overexpressed. The method can be used in vivo which
CC   enables identification of polypeptides that bind to a specific site on a
CC   DNA in the intracellular milieu. The present sequence represents the
CC   coding sequence of TG-ZFD-012 zinc finger protein which was used in
CC   the method of the invention.
XX
SQ   Sequence 69 BP; 17 A; 16 C; 20 G; 16 T; 0 other;


  Query Match             2.5%;  Score 69;  DB 22;  Length 69;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-23;
  Matches   69;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1269 tatgtttgcagggaatgtgggcgtggctttcgccagcattcacacctggtcagacacaag 1328
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 tatgtttgcagggaatgtgggcgtggctttcgccagcattcacacctggtcagacacaag 60

Qy   1329 aggacacat 1337
           |||||||||
Db     61 aggacacat 69
```

```
RESULT   7
AAS13088
ID   AAS13088 standard; DNA; 69 BP.
XX
AC   AAS13088;
XX
DT   17-DEC-2001  (first entry)
XX
DE   DNA encoding zinc finger domain TG-ZFD-046.
XX
KW   Zinc finger domain; cancer; human; ds.
XX
OS   Homo sapiens.
XX
PN   WO200160970-A2.
XX
PD   23-AUG-2001.
XX
PF   17-FEB-2001; 2001WO-KR00244.
XX
PR   18-FEB-2000; 2000KR-0007730.
XX
PA   (TOOL-) TOOLGEN INC.
XX
PI   Kim J,  Kwon Y,  Kim H,  Ryu EH,  Hwang MS;
XX
DR   WPI; 2001-557644/62.
DR   P-PSDB; AAU08558.
XX
PT   Identifying a zinc finger domain for e.g. designing new polypeptides
PT   that bind to a specific site on a DNA, comprises expressing hybrid
PT   nucleic acids with a test zinc finger domain in cells -
XX
PS   Example 54; Page 69; 147pp; English.
XX
CC   The invention relates to a method of identifying a zinc finger domain
CC   that recognises a target site on a DNA. The method comprises expressing
CC   hybrid nucleic acids with a test zinc finger domain in cells containing a
CC   reporter construct, where the reporter gene is expressed above a given
CC   level when a transcription factor recognises a recruitment and a target
CC   site of a promoter, and not only the recruitment site of the promoter.
CG   The method is used to: (a) identify a zinc finger domain that recognises
CC   a target site on a DNA; (b) determine whether a test zinc finger domain
CC   recognises a target site on a promoter; (c) generate a nucleic acid that
CC   encodes a chimeric zinc finger protein; and (d) identify DNA sequences
CC   recognised by zinc finger domains. The method can be used to design novel
CG   polypeptides that bind to a specific site on a DNA. The method can
CC   facilitate the customised generation of new polypeptides that can
CC   regulate the expression a selected target e.g. a gene required by a
CC   pathogen can be repressed, a gene required for cancerous growth can be
CC   repressed, or a gene poorly expressed or encoding a mutated protein can
CC   be activated and overexpressed. The method can be used in vivo which
CC   enables identification of polypeptides that bind to a specific site on a
CC   DNA in the intracellular milieu. The present sequence represents the
CC   coding sequence of TG-ZFD-046 zinc finger protein which was used in
CC   the method of the invention.
XX
SQ   Sequence 69 BP; 21 A; 14 C; 18 G; 16 T; 0 other;


  Query Match             2.5%;  Score 69;  DB 22;  Length 69;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-23;
  Matches   69;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    2019 tatatttgcagaaagtgtggacggggctttagtcggaagtccaaccttatcagacatcag 2078
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 tatatttgcagaaagtgtggacggggctttagtcggaagtccaaccttatcagacatcag 60

Qy    2079 aggacacac 2087
           |||||||||
Db      61 aggacacac 69
```

```
RESULT   8
AAS68257
ID   AAS68257 standard; cDNA; 821 BP.
XX
AC   AAS68257;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #4061.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG04070.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity -
XX
PS   Claim 1; SEQ ID No 4061; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human
CC   diagnostic coding sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 821 BP; 174 A; 234 C; 215 G; 198 T; 0 other;


  Query Match             1.8%;  Score 49;  DB 23;  Length 821;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-13;
  Matches   49;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     722 acaacagccagcacagtccaaggaagacaacacagtggtggatataggg 770
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db     188 acaacagccagcacagtccaaggaagacaacacagtggtggatataggg 236


RESULT   9
```

```
AAH33832
ID   AAH33832 standard; cDNA; 800 BP.
XX
AC   AAH33832;
XX
DT   03-SEP-2001  (first entry)
XX
DE   Human colon cancer antigen encoding cDNA SEQ ID NO:888.
XX
KW   Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW   colorectal carcinoma; ss.
XX
OS   Homo sapiens.
XX
PN   WO200122920-A2.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26524.
XX
PR   29-SEP-1999;   99US-0157137.
PR   03-NOV-1999;   99US-0163280.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Ruben SM,  Barash SC,  Birse CE,  Rosen CA;
XX
DR   WPI; 2001-235357/24.
DR   P-PSDB; AAG74401.
XX
PT   Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT   useful for preventing, diagnosing and/or treating colorectal cancers -
XX
PS   Claim 1; Page 2813-2814; 9803pp; English.
XX
CC   AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC   cancer-associated nucleic acid molecules (N) and proteins (P), where
CC   the proteins are collectively known as colon cancer antigens. The colon
CC   cancer antigens have cytostatic activity and can be used in gene
CC   therapy and vaccine production. N and P may be used in the prevention,
CC   diagnosis and treatment of diseases associated with inappropriate P
CC   expression. For example, N and P may be used to treat disorders
CC   associated with decreased expression by rectifying mutations or deletions
CC   in a patient's genome that affect the activity of P by expressing
CC   inactive proteins or to supplement the patients own production of P.
CC   Additionally, N may be used to produce the colon cancer-associated Ps,
CC   by inserting the nucleic acids into a host cell and culturing the cell
CC   to express the proteins. N and P can be used in the prevention, diagnosis
CC   and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC   and AAB77789 represent sequences used in the exemplification of the
CC   present invention.
CC   N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC   missing at time of publication, meaning no sequences are present for
CC   SEQ ID NO:1027 to 1052, 7921 and 7922.
XX
SQ   Sequence 800 BP; 240 A; 173 C; 219 G; 166 T; 2 other;


  Query Match             1.0%;  Score 29;  DB 22;  Length 800;
  Best Local Similarity   100.0%;  Pred. No. 0.00079;
  Matches   29;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1659 caccagaggacacactcaggggagaagcc 1687
          ||||||||||||||||||||||||||||||
Db    413 caccagaggacacactcaggggagaagcc 441


RESULT  10
AAS90429
ID   AAS90429 standard; cDNA; 2349 BP.
XX
AC   AAS90429;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #26233.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG26242.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity  -
XX
PS   Claim 1; SEQ ID No 26233; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human
CC   diagnostic coding sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 2349 BP; 664 A; 547 C; 608 G; 530 T; 0 other;


  Query Match             1.0%;  Score 29;  DB 23;  Length 2349;
  Best Local Similarity   100.0%;  Pred. No. 0.00079;
  Matches   29;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1659 caccagaggacacactcaggggagaagcc 1687
          ||||||||||||||||||||||||||||||
Db   1645 caccagaggacacactcaggggagaagcc 1673


RESULT  11
AAI58664
ID   AAI58664 standard; cDNA; 3582 BP.
XX
AC   AAI58664;
XX
DT   22-OCT-2001  (first entry)
```

XX
DE Human polynucleotide SEQ ID NO 867.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI; 2001-442253/47.
DR P-PSDB; AAM39508.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
PS Claim 1; SEQ ID NO 867; 10078pp; English.
XX
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 3582 BP; 1024 A; 826 C; 910 G; 821 T; 1 other;

Query Match 1.0%; Score 29; DB 22; Length 3582;
Best Local Similarity 100.0%; Pred. No. 0.00078;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1659 caccagaggacacactcaggggagaagcc 1687
||||||||||||||||||||||||||||||
Db 1487 caccagaggacacactcaggggagaagcc 1515

RESULT 12
AAI60450/c
ID AAI60450 standard; cDNA; 3582 BP.
XX
AC AAI60450;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 4439.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI; 2001-442253/47.
DR P-PSDB; AAM41294.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
PS Claim 1; SEQ ID NO 4439; 10078pp; English.
XX
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 3582 BP; 821 A; 908 C; 826 G; 1026 T; 1 other;

Query Match 1.0%; Score 29; DB 22; Length 3582;
Best Local Similarity 100.0%; Pred. No. 0.00078;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1659 caccagaggacacactcaggggagaagcc 1687
||||||||||||||||||||||||||||||
Db 2096 CACCAGAGGACACACTCAGGGGAGAAGCC 2068

```
RESULT  13
AAL00216
ID   AAL00216 standard; cDNA; 159 BP.
XX
AC   AAL00216;
XX
DT   21-NOV-2001  (first entry)
XX
DE   Human reproductive system related antigen cDNA SEQ ID NO: 217.
XX
KW   Human; reproductive system related antigen; reproductive system disorder;
KW   cancer; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   WO200155320-A2.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01339.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
PR   08-SEP-2000; 2000US-0232081.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232397.
PR   14-SEP-2000; 2000US-0232398.
PR   14-SEP-2000; 2000US-0232399.
PR   14-SEP-2000; 2000US-0232400.
PR   14-SEP-2000; 2000US-0232401.
PR   14-SEP-2000; 2000US-0233063.
PR   14-SEP-2000; 2000US-0233064.
PR   14-SEP-2000; 2000US-0233065.
PR   21-SEP-2000; 2000US-0234223.
PR   21-SEP-2000; 2000US-0234274.
PR   25-SEP-2000; 2000US-0234997.
PR   25-SEP-2000; 2000US-0234998.
PR   26-SEP-2000; 2000US-0235484.
PR   27-SEP-2000; 2000US-0235834.
PR   27-SEP-2000; 2000US-0235836.
PR   29-SEP-2000; 2000US-0236327.
PR   29-SEP-2000; 2000US-0236367.
PR   29-SEP-2000; 2000US-0236368.
PR   29-SEP-2000; 2000US-0236369.
PR   29-SEP-2000; 2000US-0236370.
PR   02-OCT-2000; 2000US-0236802.
PR   02-OCT-2000; 2000US-0237037.
PR   02-OCT-2000; 2000US-0237038.
PR   02-OCT-2000; 2000US-0237039.
PR   02-OCT-2000; 2000US-0237040.
PR   13-OCT-2000; 2000US-0239935.
PR   13-OCT-2000; 2000US-0239937.
PR   20-OCT-2000; 2000US-0240960.
PR   20-OCT-2000; 2000US-0241221.
PR   20-OCT-2000; 2000US-0241785.
PR   20-OCT-2000; 2000US-0241786.
PR   20-OCT-2000; 2000US-0241787.
PR   20-OCT-2000; 2000US-0241808.
PR   20-OCT-2000; 2000US-0241809.
PR   20-OCT-2000; 2000US-0241826.
PR   01-NOV-2000; 2000US-0244617.
PR   08-NOV-2000; 2000US-0246474.
PR   08-NOV-2000; 2000US-0246475.
PR   08-NOV-2000; 2000US-0246476.
PR   08-NOV-2000; 2000US-0246477.
PR   08-NOV-2000; 2000US-0246478.
PR   08-NOV-2000; 2000US-0246523.
PR   08-NOV-2000; 2000US-0246524.
PR   08-NOV-2000; 2000US-0246525.
PR   08-NOV-2000; 2000US-0246526.
PR   08-NOV-2000; 2000US-0246527.
PR   08-NOV-2000; 2000US-0246528.
PR   08-NOV-2000; 2000US-0246532.
PR   08-NOV-2000; 2000US-0246609.
PR   08-NOV-2000; 2000US-0246610.
PR   08-NOV-2000; 2000US-0246611.
PR   08-NOV-2000; 2000US-0246613.
PR   17-NOV-2000; 2000US-0249207.
PR   17-NOV-2000; 2000US-0249208.
PR   17-NOV-2000; 2000US-0249209.
PR   17-NOV-2000; 2000US-0249210.
PR   17-NOV-2000; 2000US-0249211.
PR   17-NOV-2000; 2000US-0249212.
PR   17-NOV-2000; 2000US-0249213.
PR   17-NOV-2000; 2000US-0249214.
PR   17-NOV-2000; 2000US-0249215.
PR   17-NOV-2000; 2000US-0249216.
PR   17-NOV-2000; 2000US-0249217.
PR   17-NOV-2000; 2000US-0249218.
PR   17-NOV-2000; 2000US-0249244.
PR   17-NOV-2000; 2000US-0249245.
PR   17-NOV-2000; 2000US-0249264.
PR   17-NOV-2000; 2000US-0249265.
PR   17-NOV-2000; 2000US-0249297.
PR   17-NOV-2000; 2000US-0249299.
PR   17-NOV-2000; 2000US-0249300.
PR   01-DEC-2000; 2000US-0250160.
```

PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-465570/50.
DR P-PSDB; AAM94246.
XX
PT Isolated nucleic acid molecule encoding a reproductive system antigen
PT is used in preventing, treating or ameliorating a medical condition -
XX
PS Claim 1; SEQ ID NO 217; 1297pp + Sequence Listing; English.
XX
CC The present invention provides the protein and coding sequences of a
CC number of human reproductive system related antigens. These can be used
CC in the prevention and treatment of reproductive system disorders,
CC including cancer. The present sequence is a coding sequence of the
CC invention.
XX
SQ Sequence 159 BP; 47 A; 33 C; 47 G; 30 T; 2 other;

Query Match 1.0%; Score 28; DB 22; Length 159;
Best Local Similarity 100.0%; Pred. No. 0.0025;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1659 caccagaggacacactcaggggagaagc 1686
||||||||||||||||||||||||||||
Db 18 caccagaggacacactcaggggagaagc 45

RESULT 14
AAS25949
ID AAS25949 standard; cDNA; 1183 BP.
XX
AC AAS25949;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human cDNA encoding a novel secreted protein, Seq ID 128.
XX
KW Human; immunosuppressive; antiarthritic; ss; antirheumatic;
KW cytostatic; cardiant; vasotropic; cerebroprotective; nootropic;
KW neuroprotective; antibacterial; virucide; fungicide; opthalmalogical;
KW vulnerary; secreted protein; rheumatoid arthritis;
KW hyperproliferative disorder; cardiovascular disorder; cardiac arrest;
KW cerebrovascular disorder; cerebral ischaemia; angiogenesis;
KW nervous system disorder; Alzheimer's disease; infection; ocular disorder;
KW corneal infection; wound healing; epithelial cell proliferation;
KW skin ageing; food additive; preservative; antiproliferative.
XX
OS Homo sapiens.
XX
PN WO200155322-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01341.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.

PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-488783/53.
DR P-PSDB; AAU15962.
XX
PT New nucleic acid molecules encoding 461 human secreted proteins for
PT diagnosing, preventing, treating or ameliorating medical conditions and
PT used as food additives or preservatives -
XX
PS Claim 1; SEQ ID No 128; 980pp; English.
XX
CC The invention relates to isolated nucleic acid molecules and their
CC encoded secreted proteins. The nucleic acids and proteins are used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They
CC are also used in diagnosing a pathological condition or susceptibility
CC to a pathological condition. Antibodies to the proteins can also
CC be used in alleviating symptoms associated with the disorders and in
CC diagnostic immunoassays e.g. radioimmunoassays or enzyme linked
CC immunosorbant assays (ELISA). Disorders which are diagnosed or treated
CC include autoimmune diseases e.g. rheumatoid arthritis,
CC hyperproliferative disorders e.g. neoplasms of the breast or liver,
CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
CC e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g.
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi
CC and ocular disorders e.g. corneal infection, and many other
CC disorders listed in the specification. The polypeptides can also
CC be used to aid wound healing and epithelial cell proliferation, to
CC prevent skin aging due to sunburn, to maintain organs before
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The polypeptides can also be used
CC as a food additive or preservative to increase or decrease storage
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
CC minerals, cofactors and other nutritional components. The present
CC sequence encodes a novel secreted protein of the invention.

Query Match 1.0%; Score 27; DB 22; Length 1183;
Best Local Similarity 100.0%; Pred. No. 0.0077;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2149 tcatcagacaccagaggacacacacag 2175
||||||||||||||||||||||||||||
Db 961 tcatcagacaccagaggacacacacag 987

RESULT 15
AAS26635
ID AAS26635 standard; DNA; 2717 BP.
XX
AC AAS26635;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human genomic DNA encoding partial novel secreted protein, Seq ID 1609.
XX
KW Human; immunosuppressive; antiarthritic; ds; antirheumatic;
KW cytostatic; cardiant; vasotropic; cerebroprotective; nootropic;
KW neuroprotective; antibacterial; virucide; fungicide; opthalmalogical;
KW vulnerary; secreted protein; rheumatoid arthritis;
KW hyperproliferative disorder; cardiovascular disorder; cardiac arrest;
KW cerebrovascular disorder; cerebral ischaemia; angiogenesis;
KW nervous system disorder; Alzheimer's disease; infection; ocular disorder;
KW corneal infection; wound healing; epithelial cell proliferation;
KW skin ageing; food additive; preservative; antiproliferative.
XX
OS Homo sapiens.
XX
PN WO200155322-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01341.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.

PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-488783/53.
XX
PT New nucleic acid molecules encoding 461 human secreted proteins for
PT diagnosing, preventing, treating or ameliorating medical conditions and
PT used as food additives or preservatives -
XX
PS Disclosure; SEQ ID No 1609; 980pp; English.
XX
CC The invention relates to isolated nucleic acid molecules and their
CC encoded secreted proteins. The nucleic acids and proteins are used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They
CC are also used in diagnosing a pathological condition or susceptibility
CC to a pathological condition. Antibodies to the proteins can also

CC be used in alleviating symptoms associated with the disorders and in
CC diagnostic immunoassays e.g. radioimmunoassays or enzyme linked
CC immunosorbant assays (ELISA). Disorders which are diagnosed or treated
CC include autoimmune diseases e.g. rheumatoid arthritis,
CC hyperproliferative disorders e.g. neoplasms of the breast or liver,
CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
CC e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g.
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi
CC and ocular disorders e.g. corneal infection, and many other
CC disorders listed in the specification. The polypeptides can also
CC be used to aid wound healing and epithelial cell proliferation, to
CC prevent skin aging due to sunburn, to maintain organs before
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The polypeptides can also be used
CC as a food additive or preservative to increase or decrease storage
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
CC minerals, cofactors and other nutritional components. The present
CC sequence is a genomic DNA encoding a partial novel secreted protein of
CC the invention.

Query Match 1.0%; Score 27; DB 22; Length 2717;
Best Local Similarity 100.0%; Pred. No. 0.0076;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2149 tcatcagacaccagaggacacacacag 2175
||||||||||||||||||||||||||||
Db 962 tcatcagacaccagaggacacacacag 988

Search completed: May 17, 2002, 09:32:07
Job time: 67704 sec

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: May 16, 2002, 13:38:43 ; Search time 3793.77 Seconds
(without alignments)
9861.842 Million cell updates/sec

Title: US-09-898-556A-3
Perfect score: 2772
Sequence: 1 caggcgcgttaagctggttg.....tttctacccatcctcaccct 2772

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 13736207 seqs, 6748477542 residues

Word size : 0

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	721	26.0	1082	10	BM450709	BM450709 AGENCOURT
2	677	24.4	1035	10	BM470740	BM470740 AGENCOURT
3	585	21.1	1071	10	BE741389	BE741389 601594255
4	581	21.0	590	9	AW245709	AW245709 2823003.5
5	545	19.7	1009	10	BM449478	BM449478 AGENCOURT
c 6	510	18.4	563	9	AI669533	AI669533 wb88h09.x
7	509	18.4	535	9	AL047439	AL047439 DKFZp586I
c 8	492	17.7	558	9	AI802142	AI802142 tx30c05.x
c 9	482	17.4	537	9	BE220142	BE220142 hv67f07.x
10	471	17.0	624	9	BE178716	BE178716 PM4-HT060
c 11	458	16.5	512	10	BF111754	BF111754 7135a10.x
c 12	454	16.4	505	10	BE463501	BE463501 hw24d02.x
13	449	16.2	486	9	AW401692	AW401692 UI-HF-BK0
14	419	15.1	486	9	AV720986	AV720986 AV720986
c 15	414	14.9	613	9	AW964787	AW964787 EST376980
16	413	14.9	697	10	BE391140	BE391140 601286705
17	405	14.6	673	10	BG928720	BG928720 HNC72-1-B
c 18	402	14.5	453	9	AW513018	AW513018 xt76f03.x
c 19	399	14.4	481	9	AI276016	AI276016 qw08h02.x
20	395	14.2	451	9	AA450184	AA450184 zx42e09.r
c 21	383	13.8	435	9	AI081665	AI081665 ou63e06.s
c 22	383	13.8	449	9	AI750087	AI750087 at35f07.x
c 23	378	13.6	455	10	BE501382	BE501382 hw31a02.x
24	378	13.6	728	10	BE746403	BE746403 601579525
c 25	372	13.4	424	9	AI016684	AI016684 ou96d05.x
c 26	372	13.4	428	9	AI440109	AI440109 ti56h06.x
c 27	372	13.4	458	9	AI953623	AI953623 wq23a03.x
c 28	366	13.2	501	9	AW140124	AW140124 UI-H-BI1-
c 29	363	13.1	426	9	AI828662	AI828662 tr06f04.x
c 30	355	12.8	486	9	AI650983	AI650983 wa96d07.x
c 31	351	12.7	473	10	BF476173	BF476173 naa29c11.
32	350	12.6	385	9	AA218658	AA218658 zq96g04.r
c 33	350	12.6	500	9	AI986444	AI986444 wz65d09.x
34	345	12.4	551	10	BE395838	BE395838 601310077
c 35	344	12.4	508	9	AI860707	AI860707 wl15d12.x
c 36	343	12.4	399	9	AI289399	AI289399 qw32e05.x
c 37	340	12.3	392	9	AA218659	AA218659 zq96g04.s
c 38	340	12.3	395	9	AI361705	AI361705 qz18f04.x
39	337	12.2	513	9	AW958936	AW958936 EST371006
40	333	12.0	794	10	BI256434	BI256434 602974454
c 41	332	12.0	383	9	AA450118	AA450118 zx42e09.s
c 42	318	11.5	378	9	AA484745	AA484745 ne81h06.s
c 43	315	11.4	369	9	AI159962	AI159962 qb80f10.x
c 44	315	11.4	818	9	AL566746	AL566746 AL566746
45	312	11.3	501	10	R50617	R50617 yj57b01.r1

ALIGNMENTS

RESULT 1
BM450709
LOCUS BM450709 1082 bp mRNA linear EST 05-FEB-2002
DEFINITION AGENCOURT_6394734 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5494415 5', mRNA sequence.
ACCESSION BM450709
VERSION BM450709.1 GI:18499749
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1082)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12119 row: k column: 24
High quality sequence stop: 585.
FEATURES Location/Qualifiers
source 1. .1082
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5494415"
/clone_lib="NIH_MGC_67"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.75 kb. Library constructed by Life Technologies."
BASE COUNT 298 a 261 c 292 g 226 t 5 others

```
ORIGIN


  Query Match             26.0%;  Score 721;  DB 10;  Length 1082;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  721;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     849 atcaaatatgaagagtttgggccaggctttatcaaggagtcaaacctccttagcctccag 908
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2 ATCAAATATGAAGAGTTTGGGCCAGGCTTTATCAAGGAGTCAAACCTCCTTAGCCTCCAG 61

Qy     909 aagacacaaactggggagacaccttacatgtacactgagtggggagacagctttggcagt 968
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      62 AAGACACAAACTGGGGAGACACCTTACATGTACACTGAGTGGGGAGACAGCTTTGGCAGT 121

Qy     969 atgtcagtcctcatcaaaaacccaaggacacactctgggggaaagccttatgtgtgcagg 1028
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     122 ATGTCAGTCCTCATCAAAAACCCAAGGACACACTCTGGGGGAAAGCCTTATGTGTGCAGG 181

Qy    1029 gaatgtgggcgaggctttacgtggaagtcaaacctgatcacacatcagaggacacactca 1088
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     182 GAATGTGGGCGAGGCTTTACGTGGAAGTCAAACCTGATCACACATCAGAGGACACACTCA 241

Qy    1089 ggggagaaaccttatgtgtgcaaggattgtggacgaggctttacttggaagtcgaacctc 1148
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     242 GGGGAGAAACCTTATGTGTGCAAGGATTGTGGACGAGGCTTTACTTGGAAGTCGAACCTC 301

Qy    1149 tttacacatcagcggacacactcagggctcaagccttatgtgtgcaaggaatgtgggcag 1208
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     302 TTTACACATCAGCGGACACACTCAGGGCTCAAGCCTTATGTGTGCAAGGAATGTGGGCAG 361

Qy    1209 agctttagcctgaagtcaaacctcattacccaccagagggcgcacactggggagaagcct 1268
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     362 AGCTTTAGCCTGAAGTCAAACCTCATTACCCACCAGAGGGCGCACACTGGGGAGAAGCCT 421

Qy    1269 tatgtttgcagggaatgtgggcgtggctttcgccagcattcacacctggtcagacacaag 1328
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     422 TATGTTTGCAGGGAATGTGGGCGTGGCTTTCGCCAGCATTCACACCTGGTCAGACACAAG 481

Qy    1329 aggacacattcaggagagaagccttacatttgcagggagtgtgagcaaggctttagccag 1388
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     482 AGGACACATTCAGGAGAGAAGCCTTACATTTGCAGGGAGTGTGAGCAAGGCTTTAGCCAG 541

Qy    1389 aagtcacacctcatcagacacttaaggacacacacaggagagaagccttatgtatgcaca 1448
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     542 AAGTCACACCTCATCAGACACTTAAGGACACACACAGGAGAGAAGCCTTATGTATGCACA 601

Qy    1449 gaatgtgggcgtcactttagctggaaatcaaacctcaaaacacaccagaggacacactca 1508
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     602 GAATGTGGGCGTCACTTTAGCTGGAAATCAAACCTCAAAACACACCAGAGGACACACTCA 661

Qy    1509 ggggttaaaccttatgtctgcctggagtgcgggcagtgctttagcctgaagtcaaacctt 1568
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     662 GGGGTTAAACCTTATGTCTGCCTGGAGTGCGGGCAGTGCTTTAGCCTGAAGTCAAACCTT 721

Qy    1569 a 1569
           |
Db     722 A 722


RESULT   2
BM470740
LOCUS       BM470740                1035 bp    mRNA    linear   EST 05-FEB-2002
DEFINITION  AGENCOURT_6475358 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5579278
            5', mRNA sequence.
ACCESSION   BM470740
VERSION     BM470740.1  GI:18519782
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1035)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12335  row: k  column: 23
            High quality sequence stop: 653.
FEATURES             Location/Qualifiers
     source          1. .1035
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5579278"
                     /clone_lib="NIH_MGC_67"
                     /tissue_type="retinoblastoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.75 kb.   Library constructed by Life
                     Technologies."
BASE COUNT      259 a    249 c    268 g    255 t      4 others
ORIGIN


  Query Match             24.4%;  Score 677;  DB 10;  Length 1035;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  677;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    1655 cacgcaccagaggacacactcaggggagaagccatttgtatgtgctgagtgtggacgagg 1714
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25 CACGCACCAGAGGACACACTCAGGGGAGAAGCCATTTGTATGTGCTGAGTGTGGACGAGG 84

Qy    1715 ctttaatgataagtccaccctcatttcacaccagaggacacattcaggggaaaagccttt 1774
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      85 CTTTAATGATAAGTCCACCCTCATTTCACACCAGAGGACACATTCAGGGGAAAAGCCTTT 144

Qy    1775 tatgtgcagggagtgtggcagaaggtttcggcagaagcctaacctgtttaggcacaagag 1834
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     145 TATGTGCAGGGAGTGTGGCAGAAGGTTTCGGCAGAAGCCTAACCTGTTTAGGCACAAGAG 204

Qy    1835 ggcacactcaggtgccttttgtgtgcagggagtgtgggcaaggcttttgtgctaagttaac 1894
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     205 GGCACACTCAGGTGCCTTTGTGTGCAGGGAGTGTGGGCAAGGCTTTTGTGCTAAGTTAAC 264

Qy    1895 tctcattaaacaccagagagcacacgcagggggggaagcctcatgtgtgcagggagtgtgg 1954
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     265 TCTCATTAAACACCAGAGAGCACACGCAGGGGGGAAGCCTCATGTGTGCAGGGAGTGTGG 324

Qy    1955 gcaaggctttagccggcagtcacacctcattagacaccagaggacacattcaggagagaa 2014
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     325 GCAAGGCTTTAGCCGGCAGTCACACCTCATTAGACACCAGAGGACACATTCAGGAGAGAA 384

Qy    2015 gccttatatttgcagaaagtgtggacggggctttagtcggaagtccaaccttatcagaca 2074
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     385 GCCTTATATTTGCAGAAAGTGTGGACGGGGCTTTAGTCGGAAGTCCAACCTTATCAGACA 444

Qy    2075 tcagaggacacactcaggatagaaactttatgtgtatagggaatgtggtacagcctttag 2134
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     445 TCAGAGGACACACTCAGGATAGAAACTTTATGTGTATAGGGAATGTGGTACAGCCTTTAG 504

Qy    2135 ccaggagtcatacttcatcagacaccagaggacacacacagtgctgtggctttttcagcc 2194
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     505 CCAGGAGTCATACTTCATCAGACACCAGAGGACACACACAGTGCTGTGGCTTTTTCAGCC 564

Qy    2195 attgctagataccaaagtggagacattctgtgtgtgattatgcatgagactgtactggta 2254
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     565 ATTGCTAGATACCAAAGTGGAGACATTCTGTGTGTGATTATGCATGAGACTGTACTGGTA 624
```

```
Qy   2255 agacttgtatctccatccacctgaaggagaattgctggctcattttcaggagccctgccc 2314
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    625 AGACTTGTATCTCCATCCACCTGAAGGAGAATTGCTGGCTCATTTTCAGGAGCCCTGCCC 684

Qy   2315 ttcctcactgtggatgg 2331
          |||||||||||||||||
Db    685 TTCCTCACTGTGGATGG 701


RESULT   3
BE741389
LOCUS       BE741389                1071 bp    mRNA    linear   EST 15-SEP-2000
DEFINITION  601594255F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3948468 5',
            mRNA sequence.
ACCESSION   BE741389
VERSION     BE741389.1  GI:10155381
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1071)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM811  row: m  column: 13
            High quality sequence stop: 783.
FEATURES             Location/Qualifiers
     source          1. .1071
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3948468"
                     /clone_lib="NIH_MGC_9"
                     /tissue_type="adenocarcinoma cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: ovary; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Size-selected >500bp for average
                     insert size 1.8kb.  Library constructed by Ling Hong in
                     the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT      287 a    233 c    293 g    257 t      1 others
ORIGIN


  Query Match           21.1%;  Score 585;  DB 10;  Length 1071;
  Best Local Similarity  99.8%;  Pred. No. 7.5e-297;
  Matches  635;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy   1678 gggagaagccatttgtatgtgctgagtgtggacgaggctttaatgataagtccaccctca 1737
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     36 GGGAGAAGCCATTTGTATGTGCTGAGTGTGGACGAGGCTTTAATGATAAGTCCACCCTCA 95

Qy   1738 tttcacaccagaggacacattcaggggaaaagccttttatgtgcagggagtgtggcagaa 1797
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     96 TTTCACACCAGAGGACACATTCAGGGGAAAAGCCTTTTATGTGCAGGGAGTGTGGCAGAA 155

Qy   1798 ggtttcggcagaagcctaacctgtttaggcacaagagggcacactcaggtgcctttgtgt 1857
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    156 GGTTTCGGCAGAAGCCTAACCTGTTTAGGCACAAGAGGGCACACTCAGGTGCCTTTGTGT 215

Qy   1858 gcagggagtgtgggcaaggcttttgtgctaagttaactctcattaaacaccagagagcac 1917
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    216 GCAGGGAGTGTGGGCAAGGCTTTTGTGCTAAGTTAACTCTCATTAAACACCAGAGAGCAC 275

Qy   1918 acgcaggggggaagcctcatgtgtgcagggagtgtgggcaaggctttagccggcagtcac 1977
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    276 ACGCAGGGGGGAAGCCTCATGTGTGCAGGGAGTGTGGGCAAGGCTTTAGCCGGCAGTCAC 335

Qy   1978 acctcattagacaccagaggacacattcaggagagaagccttatatttgcagaaagtgtg 2037
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    336 ACCTCATTAGACACCAGAGGACACATTCAGGAGAGAAGCCTTATATTTGCAGAAAGTGTG 395

Qy   2038 gacggggctttagtcggaagtccaaccttatcagacatcagaggacacactcaggataga 2097
          |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db    396 GACGGGGCTTTAGTCGGAAGTCCAACCTTATCAGACATCACAGGACACACTCAGGATAGA 455

Qy   2098 aactttatgtgtatagggaatgtggtacagcctttagccaggagtcatacttcatcagac 2157
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    456 AACTTTATGTGTATAGGGAATGTGGTACAGCCTTTAGCCAGGAGTCATACTTCATCAGAC 515

Qy   2158 accagaggacacacacagtgctgtggctttttcagccattgctagataccaaagtggaga 2217
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    516 ACCAGAGGACACACACAGTGCTGTGGCTTTTTCAGCCATTGCTAGATACCAAAGTGGAGA 575

Qy   2218 cattctgtgtgtgattatgcatgagactgtactggtaagacttgtatctccatccacctg 2277
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    576 CATTCTGTGTGTGATTATGCATGAGACTGTACTGGTAAGACTTGTATCTCCATCCACCTG 635

Qy   2278 aaggagaattgctggctcattttcaggagccctgcc 2313
          |||||||||||||||||||||||||||||||||||
Db    636 AAGGAGAATTGCTGGCTCATTTTCAGGAGCCCTGCC 671


RESULT   4
AW245709
LOCUS       AW245709                590 bp    mRNA    linear   EST 07-JAN-2000
DEFINITION  2823003.5prime NIH_MGC_7 Homo sapiens cDNA clone IMAGE:2823003 5',
            mRNA sequence.
ACCESSION   AW245709
VERSION     AW245709.1  GI:6588702
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 590)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Other_ESTs: 2823003.3prime
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP cDNA Library Preparation: Ling
            Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E.
            Consortium (LLNL) DNA Sequencing by: Berkeley MGC sequencing
            project Clone distribution: MGC clone distribution information can
            be found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html Base Calling / Quality
            Scores: PHRED from University of Washington Genome Center. Vector
            Trimming: cross_match from University of Washington Genome Center
            PHRAP suite. Poly-T Identification: patMatch.pl from Berkeley
            Drosophila Genome Project. University of Washington Genome Center:
            http://www.genome.washington.edu
            Plate: LLCM10  row: O  column: 4
            High quality sequence stop: 582.
FEATURES             Location/Qualifiers
     source          1. .590
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2823003"
                     /clone_lib="NIH_MGC_7"
                     /tissue_type="small cell carcinoma"
                     /cell_line="MGC3"
                     /lab_host="DH10B (phage-resistant)"
```

/note="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 192 a 127 c 161 g 110 t
ORIGIN

Query Match 21.0%; Score 581; DB 9; Length 590;
Best Local Similarity 100.0%; Pred. No. 8.7e-295;
Matches 581; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    544 gggcaggaaatcctctccacctgggaaaacactatccagaagatcagaaacaacagcagg 603
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     10 GGGCAGGAAATCCTCTCCACCTGGGAAAACACTATCCAGAAGATCAGAAACAACAGCAGG 69

Qy    604 atccattctgctttagtggcaaagcagaatggattcaagagggagaagactccagactcc 663
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     70 ATCCATTCTGCTTTAGTGGCAAAGCAGAATGGATTCAAGAGGGAGAAGACTCCAGACTCC 129

Qy    664 tgtttgggagagtaagcaaaaatggcacttcaaaggcactttccagcccacctgaagaac 723
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    130 TGTTTGGGAGAGTAAGCAAAAATGGCACTTCAAAGGCACTTTCCAGCCCACCTGAAGAAC 189

Qy    724 aacagccagcacagtccaaggaagacaacacagtggtggatatagggtccagccctgaac 783
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    190 AACAGCCAGCACAGTCCAAGGAAGACAACACAGTGGTGGATATAGGGTCCAGCCCTGAAC 249

Qy    784 ggagggcagatctagaggaaacagacaaagtattgcatggtttagaagtctcaggatttg 843
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    250 GGAGGGCAGATCTAGAGGAAACAGACAAAGTATTGCATGGTTTAGAAGTCTCAGGATTTG 309

Qy    844 gagaaatcaaatatgaagagtttgggccaggctttatcaaggagtcaaacctccttagcc 903
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    310 GAGAAATCAAATATGAAGAGTTTGGGCCAGGCTTTATCAAGGAGTCAAACCTCCTTAGCC 369

Qy    904 tccagaagacacaaactggggagacaccttacatgtacactgagtggggagacagctttg 963
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    370 TCCAGAAGACACAAACTGGGGAGACACCTTACATGTACACTGAGTGGGGAGACAGCTTTG 429

Qy    964 gcagtatgtcagtcctcatcaaaaacccaaggacacactctgggggaaagccttatgtgt 1023
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    430 GCAGTATGTCAGTCCTCATCAAAAACCCAAGGACACACTCTGGGGGAAAGCCTTATGTGT 489

Qy   1024 gcagggaatgtgggcgaggctttacgtggaagtcaaacctgatcacacatcagaggacac 1083
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    490 GCAGGGAATGTGGGCGAGGCTTTACGTGGAAGTCAAACCTGATCACACATCAGAGGACAC 549

Qy   1084 actcaggggagaaaccttatgtgtgcaaggattgtggacga 1124
          |||||||||||||||||||||||||||||||||||||||
Db    550 ACTCAGGGGAGAAACCTTATGTGTGCAAGGATTGTGGACGA 590
```

RESULT 5
BM449478
LOCUS BM449478 1009 bp mRNA linear EST 05-FEB-2002
DEFINITION AGENCOURT_6401009 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5493685 5', mRNA sequence.
ACCESSION BM449478
VERSION BM449478.1 GI:18498518
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1009)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12117 row: m column: 14
High quality sequence stop: 649.
FEATURES Location/Qualifiers
source 1. .1009
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5493685"
/clone_lib="NIH_MGC_67"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.75 kb. Library constructed by Life Technologies."
BASE COUNT 289 a 235 c 259 g 224 t 2 others
ORIGIN

Query Match 19.7%; Score 545; DB 10; Length 1009;
Best Local Similarity 99.8%; Pred. No. 9.4e-276;
Matches 595; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy    206 ggcgttcgtggcattcagggatgtggctgtgtacttcacccaggaggagtggaggttgtt 265
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    196 GGCGTTCGTGGCATTCAGGGATGTGGCTGTGTACTTCACCCAGGAGGAGTGGAGGTTGTT 255

Qy    266 gagccctgctcagaggaccctgcacagggaggtgatgctggagacttataaccatctggt 325
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    256 GAGCCCTGCTCAGAGGACCCTGCACAGGGAGGTGATGCTGGAGACTTATAACCATCTGGT 315

Qy    326 ctcactggaaattccatcttctaaaccaaaaactcattgctcagctggagcgaggggaagc 385
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    316 CTCACTGGAAATTCCATCTTCTAAACCAAAACTCATTGCTCAGCTGGAGCGAGGGGAAGC 375

Qy    386 gccctggagagaggagagaaaatgtccactggacctctgtccagaatcgaagccagaaat 445
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    376 GCCCTGGAGAGAGGAGAGAAAATGTCCACTGGACCTCTGTCCAGAATCGAAGCCAGAAAT 435

Qy    446 tcaacttagtccctcctgccctctgattttctccagtcagcaagctctcagccaacatgt 505
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    436 TCAACTTAGTCCCTCCTGCCCTCTGATTTTCTCCAGTCAGCAAGCTCTCAGCCAACATGT 495

Qy    506 gtggctgagtcatctctctcagctgttttcaagtttatgggcaggaaatcctctccacct 565
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    496 GTGGCTGAGTCATCTCTCTCAGCTGTTTTCAAGTTTATGGGCAGGAAATCCTCTCCACCT 555

Qy    566 gggaaaacactatccagaagatcagaaacaacagcaggatccattctgctttagtggcaa 625
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    556 GGGAAAACACTATCCAGAAGATCAGAAACAACAGCAGGATCCATTCTGCTTTAGTGGCAA 615

Qy    626 agcagaatggattcaagagggagaagactccagactcctgtttgggagagtaagcaaaaa 685
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    616 AGCAGAATGGATTCAAGAGGGAGAAGACTCCAGACTCCTGTTTGGGAGAGTAAGCAAAAA 675

Qy    686 tggcacttcaaaggcactttccagcccacctgaagaacaacagccagcacagtccaagga 745
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    676 TGGCACTTCAAAGGCACTTTCCAGCCCACCTGAAGAACAACAGCCAGCACAGTCCAAGGA 735

Qy    746 agacaacacagtggtggatatagggtccagccctgaacggagggcagatctagagg 801
          |||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    736 AGACNACACAGTGGTGGATATAGGGTCCAGCCCTGAACGGAGGGCAGATCTAGAGG 791
```

RESULT 6

```
AI669533/c
LOCUS       AI669533                563 bp    mRNA    linear   EST 14-MAY-1999
DEFINITION  wb88h09.x1 NCI_CGAP_Pr28 Homo sapiens cDNA clone IMAGE:2312801 3',
            mRNA sequence.
ACCESSION   AI669533
VERSION     AI669533.1  GI:4834307
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 563)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: -40UP from Gibco
            High quality sequence stop: 454.
FEATURES             Location/Qualifiers
     source          1. .563
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2312801"
                     /clone_lib="NCI_CGAP_Pr28"
                     /sex="male"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia)
                     with a modified polylinker; Plasmid DNA from the
                     normalized library NCI_CGAP_Pr22 was prepared, and ss
                     circles were made in vitro. Following HAP purification,
                     this DNA was used as tracer in a subtractive hybridization
                     reaction. The driver was PCR-amplified cDNAs from a pool
                     of 5,000 clones made from the same library (cloneIDs
                     985608-986759, 1101192-1101959, and 1217928-1220615).
                     Subtraction by Bento Soares and M. Fatima Bonaldo.  "
BASE COUNT      162 a    125 c    150 g    126 t
ORIGIN
```

```
  Query Match             18.4%;  Score 510;  DB 9;  Length 563;
  Best Local Similarity   99.8%;  Pred. No. 2.5e-257;
  Matches  560;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy    2212 tggagacattctgtgtgtgattatgcatgagactgtactggtaagacttgtatctccatc 2271
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     563 TGGAGACATTCTGTGTGTGATTATGCATGAGACTGTACTGGTAAGACTTGTATCTCCATC 504

Qy    2272 cacctgaaggagaattgctggctcattttcaggagccctgcccttcctcactgtggatgg 2331
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     503 CACCTGAAGGAGAATTGCTGGCTCATTTTCAGGAGCCCTGCCCTTCCTCACTGTGGATGG 444

Qy    2332 tgggttgtggaaacccggtcaggtaatgatagtggcaggaggcagtcaaatgcccaggca 2391
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     443 TGGGTTGTGGAAACCCGGTCAGGTAATGATAGTGGCAGGAGGCAGTCAAATGCCCAGGCA 384

Qy    2392 gataggggtgggtacctggtgaaacccaaccttaaagctgaagacagtcccggctaaatc 2451
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     383 GATAGGGGTGGGTACCTGGTGAAACCCAACCTTAAAGCTGAAGACAGTCCCGGCTAAATC 324

Qy    2452 ctcatactgaattgagaacctgtcttcccatttggtgtgctttcctccgattgatcccaa 2511
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     323 CTCATACTGAATTGAGAACCTGTCTTCCCATTTGGTGTGCTTTCCTCCGATTGATCCCAA 264
```

```
Qy    2512 cccttcacctatttacgtatacctgccctttcctaattggtttttacactgctgtgccc 2571
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     263 CCCTTCACCTATTTTACGTATACCTGCCCTTTCCTAATTGGTTTTTACACTGCTGTGCCC 204

Qy    2572 acctttgagtggtgcctttgcatacttacaaatcagtcaacgtgtattcccctattctg 2631
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     203 ACCTTTTGAGTGGTGCCTTTGCATACTTACAAATCAGTCAACGTGTATTCCCCTATTCTG 144

Qy    2632 agcccataaaagacccagactcagctgcagtgaggagagaaatcaccctgctgtggaggt 2691
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db     143 AGCCCATAAAAGACCCAGACTCAGCTGCAGTGAGGAGAGAAATCACCCTGCTGTGGGGGT 84

Qy    2692 tggggaccactccctgcatcccctctccactgagagctgttcttttgctcaataaaattc 2751
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      83 TGGGGACCACTCCCTGCATCCCCTCTCCACTGAGAGCTGTTCTTTTGCTCAATAAAATTC 24

Qy    2752 ttttctacccatcctcaccct 2772
           ||||||||||||||||||||
Db      23 TTTTCTACCCATCCTCACCCT 3
```

```
RESULT   7
AL047439
LOCUS       AL047439                535 bp    mRNA    linear   EST 29-FEB-2000
DEFINITION  DKFZp586I0320_r1 586 (synonym: hutel) Homo sapiens cDNA clone
            DKFZp586I0320, mRNA sequence.
ACCESSION   AL047439
VERSION     AL047439.1  GI:4727354
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 535)
  AUTHORS   Wambutt,R., Heubner,D., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
  TITLE     EST (Wambutt, et al.)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Wambutt R
            MIPS
            Am Klopferspitz 18a D-82152 Martinsried, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
            sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
            consortium of the German Genome Project.
            No s1 sequence available.
            This clone (DKFZp586I0320) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES             Location/Qualifiers
     source          1. .535
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="DKFZp586I0320"
                     /clone_lib="586 (synonym: hutel)"
                     /tissue_type="uterus"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Vector: pSport1; Site_1: NotI; Site_2: SalI/MluI"
BASE COUNT      159 a    128 c    143 g    105 t
ORIGIN
```

```
  Query Match             18.4%;  Score 509;  DB 9;  Length 535;
  Best Local Similarity   100.0%;  Pred. No. 8.2e-257;
  Matches  509;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    1071 catcagaggacacactcaggggagaaaccttatgtgtgcaaggattgtggacgaggcttt 1130
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CATCAGAGGACACACTCAGGGGAGAAACCTTATGTGTGCAAGGATTGTGGACGAGGCTTT 60

Qy    1131 acttggaagtcgaacctctttacacatcagcggacacactcagggctcaagccttatgtg 1190
```

```
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ACTTGGAAGTCGAACCTCTTTACACATCAGCGGACACACTCAGGGCTCAAGCCTTATGTG 120

Qy    1191 tgcaaggaatgtgggcagagctttagcctgaagtcaaacctcattacccaccagagggcg 1250
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 TGCAAGGAATGTGGGCAGAGCTTTAGCCTGAAGTCAAACCTCATTACCCACCAGAGGGCG 180

Qy    1251 cacactggggagaagccttatgtttgcagggaatgtgggcgtggctttcgccagcattca 1310
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 CACACTGGGGAGAAGCCTTATGTTTGCAGGGAATGTGGGCGTGGCTTTCGCCAGCATTCA 240

Qy    1311 cacctggtcagacacaagaggacacattcaggagagaagccttacatttgcagggagtgt 1370
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CACCTGGTCAGACACAAGAGGACACATTCAGGAGAGAAGCCTTACATTTGCAGGGAGTGT 300

Qy    1371 gagcaaggctttagccagaagtcacacctcatcagacacttaaggacacacacaggagag 1430
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 GAGCAAGGCTTTAGCCAGAAGTCACACCTCATCAGACACTTAAGGACACACACAGGAGAG 360

Qy    1431 aagccttatgtatgcacagaatgtgggcgtcactttagctggaaatcaaacctcaaaaca 1490
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 AAGCCTTATGTATGCACAGAATGTGGGCGTCACTTTAGCTGGAAATCAAACCTCAAAACA 420

Qy    1491 caccagaggacacactcaggggttaaaccttatgtctgcctggagtgcgggcagtgcttt 1550
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 CACCAGAGGACACACTCAGGGGTTAAACCTTATGTCTGCCTGGAGTGCGGGCAGTGCTTT 480

Qy    1551 agcctgaagtcaaaccttaacaaacacca 1579
                |||||||||||||||||||||||||||||
Db     481 AGCCTGAAGTCAAACCTTAACAAACACCA 509


RESULT   8
AI802142/c
LOCUS       AI802142                558 bp    mRNA    linear   EST 16-DEC-1999
DEFINITION  tx30c05.x1 NCI_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:2271080 3',
            mRNA sequence.
ACCESSION   AI802142
VERSION     AI802142.1  GI:5367614
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 558)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 771   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 455.
FEATURES             Location/Qualifiers
     source          1. .558
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2271080"
                     /clone_lib="NCI_CGAP_Lu24"
                     /tissue_type="carcinoid"
                     /lab_host="DH10B"
                     /note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Plasmid DNA from the normalized
                     library NCI_CGAP_Lu5 was prepared, and ss circles were
                     made in vitro. Following HAP purification, this DNA was
                     used as tracer in a subtractive hybridization reaction.
                     The driver was PCR-amplified cDNAs from a pool of 5,000
                     clones made from the same library (cloneIDs
                     1414920-1417991 and 1520904-1522439). Subtraction by Bento
                     Soares and M. Fatima Bonaldo. "
BASE COUNT      161 a    123 c    148 g    126 t
ORIGIN


  Query Match             17.7%;  Score 492;  DB 9;  Length 558;
  Best Local Similarity   99.8%;  Pred. No. 7.7e-248;
  Matches  542;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy    2216 gacattctgtgtgtgattatgcatgagactgtactggtaagacttgtatctccatccacc 2275
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     558 GACATTCTGTGTGTGATTATGCATGAGACTGTACTGGTAAGACTTGTATCTCCATCCACC 499

Qy    2276 tgaaggagaattgctggctcattttcaggagccctgcccttcctcactgtggatggtggg 2335
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     498 TGAAGGAGAATTGCTGGCTCATTTTCAGGAGCCCTGCCCTTCCTCACTGTGGATGGTGGG 439

Qy    2336 ttgtggaaacccggtcaggtaatgatagtggcaggaggcagtcaaatgcccaggcagata 2395
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     438 TTGTGGAAACCCGGTCAGGTAATGATAGTGGCAGGAGGCAGTCAAATGCCCAGGCAGATA 379

Qy    2396 ggggtgggtacctggtgaaacccaaccttaaagctgaagacagtcccggctaaatcctca 2455
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     378 GGGGTGGGTACCTGGTGAAACCCAACCTTAAAGCTGAAGACAGTCCCGGCTAAATCCTCA 319

Qy    2456 tactgaattgagaacctgtcttcccatttggtgtgctttcctccgattgatcccaaccct 2515
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     318 TACTGAATTGAGAACCTGTCTTCCCATTTGGTGTGCTTTCCTCCGATTGATCCCAACCCT 259

Qy    2516 tcacctattttacgtatacctgccctttcctaattggtttttacactgctgtgcccacct 2575
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     258 TCACCTATTTTACGTATACCTGCCCTTTCCTAATTGGTTTTTACACTGCTGTGCCCACCT 199

Qy    2576 tttgagtggtgcctttgcatacttacaaatcagtcaacgtgtattcccctattctgagcc 2635
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     198 TTTGAGTGGTGCCTTTGCATACTTACAAATCAGTCAACGTGTATTCCCCTATTCTGAGCC 139

Qy    2636 cataaaagacccagactcagctgcagtgaggagagaaatcaccctgctgtggaggttggg 2695
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     138 CATAAAAGACCCAGACTCAGCTGCAGTGAGGAGAGAAATCACCCTGCTGTGGGGGTTGGG 79

Qy    2696 gaccactccctgcatcccctctccactgagagctgttcttttgctcaataaaattctttt 2755
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      78 GACCACTCCCTGCATCCCCTCTCCACTGAGAGCTGTTCTTTTGCTCAATAAAATTCTTTT 19

Qy    2756 cta 2758
                |||
Db      18 CTA 16


RESULT   9
BE220142/c
LOCUS       BE220142                537 bp    mRNA    linear   EST 03-JUL-2000
DEFINITION  hv67f07.x1 NCI_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:3178501 3',
            mRNA sequence.
ACCESSION   BE220142
VERSION     BE220142.1  GI:8907460
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 537)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
```

```
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL, send email to:
            info@image.llnl.gov
            Seq primer: -40UP from Gibco
            High quality sequence stop: 450.
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3178501"
                     /clone_lib="NCI_CGAP_Lu24"
                     /tissue_type="carcinoid"
                     /lab_host="DH10B"
                     /note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Plasmid DNA from the normalized
                     library NCI_CGAP_Lu5 was prepared, and ss circles were
                     made in vitro. Following HAP purification, this DNA was
                     used as tracer in a subtractive hybridization reaction.
                     The driver was PCR-amplified cDNAs from a pool of 5,000
                     clones made from the same library (cloneIDs
                     1414920-1417991 and 1520904-1522439). Subtraction by Bento
                     Soares and M. Fatima Bonaldo. "
BASE COUNT      150 a    118 c    144 g    125 t
ORIGIN


  Query Match            17.4%;  Score 482;  DB 9;  Length 537;
  Best Local Similarity   99.8%;  Pred. No. 1.4e-242;
  Matches  532;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy   2236 gcatgagactgtactggtaagacttgtatctccatccacctgaaggagaattgctggctc 2295
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    537 GCATGAGACTGTACTGGTAAGACTTGTATCTCCATCCACCTGAAGGAGAATTGCTGGCTC 478

Qy   2296 attttcaggagccctgcccttcctcactgtggatggtgggttgtggaaacccggtcaggt 2355
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    477 ATTTTCAGGAGCCCTGCCCTTCCTCACTGTGGATGGTGGGTTGTGGAAACCCGGTCAGGT 418

Qy   2356 aatgatagtggcaggaggcagtcaaatgcccaggcagataggggtgggtacctggtgaaa 2415
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    417 AATGATAGTGGCAGGAGGCAGTCAAATGCCCAGGCAGATAGGGGTGGGTACCTGGTGAAA 358

Qy   2416 cccaaccttaaagctgaagacagtcccggctaaatcctcatactgaattgagaacctgtc 2475
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    357 CCCAACCTTAAAGCTGAAGACAGTCCCGGCTAAATCCTCATACTGAATTGAGAACCTGTC 298

Qy   2476 ttcccatttggtgtgctttcctccgattgatcccaacccttcacctattttacgtatacc 2535
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    297 TTCCCATTTGGTGTGCTTTCCTCCGATTGATCCCAACCCTTCACCTATTTTACGTATACC 238

Qy   2536 tgccctttcctaattggtttttacactgctgtgcccacctttgagtggtgcctttgcat 2595
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    237 TGCCCTTTCCTAATTGGTTTTTACACTGCTGTGCCCACCTTTTGAGTGGTGCCTTTGCAT 178

Qy   2596 acttacaaatcagtcaacgtgtattcccctattctgagcccataaaagacccagactcag 2655
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    177 ACTTACAAATCAGTCAACGTGTATTCCCCTATTCTGAGCCCATAAAAGACCCAGACTCAG 118

Qy   2656 ctgcagtgaggagagaaatcaccctgctgtggaggttggggaccactccctgcatccccct 2715
           |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db    117 CTGCAGTGAGGAGAGAAATCACCCTGCTGTGGGGGTTGGGGACCACTCCCTGCATCCCCT 58

Qy   2716 ctccactgagagctgttcttttgctcaataaaattcttttctacccatcctca 2768
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     57 CTCCACTGAGAGCTGTTCTTTTGCTCAATAAAATTCTTTTCTACCCATCCTCA 5
```

```
RESULT  10
BE178716
LOCUS       BE178716                624 bp    mRNA    linear   EST 22-JUN-2000
DEFINITION  PM4-HT0606-030400-001-c10 HT0606 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BE178716
VERSION     BE178716.1  GI:8657868
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 624)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=PM4-HT0606-030
            400-001-c10&t3=2000-04-03&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 9
            High quality sequence stop: 603.
FEATURES             Location/Qualifiers
     source          1. .624
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="HT0606"
                     /dev_stage="Adult"
                     /note="Organ: head_neck; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
BASE COUNT      188 a    137 c    177 g    122 t
ORIGIN


  Query Match            17.0%;  Score 471;  DB 9;  Length 624;
  Best Local Similarity   99.7%;  Pred. No. 9.4e-237;
  Matches  571;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy    734 acagtccaaggaagacaacacagtggtggatatagggtccagccctgaacggagggcaga 793
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      9 ACAGTCCAAGGAAGACAACACAGTGGTGGATATAGGGTCCAGCCCTGAACGGAGGGCAGA 68

Qy    794 tctagaggaaacagacaaagtattgcatggtttagaagtctcaggatttggagaaatcaa 853
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     69 TCTAGAGGAAACAGACAAAGTATTGCATGGTTTAGAAGTCTCAGGATTTGGAGAAATCAA 128

Qy    854 atatgaagagtttgggccaggctttatcaaggagtcaaacctccttagcctccagaagac 913
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    129 ATATGAAGAGTTTGGGCCAGGCTTTATCAAGGAGTCAAACCTCCTTAGCCTCCAGAAGAC 188

Qy    914 acaaactggggagacaccttacatgtacactgagtggggagacagctttggcagtatgtc 973
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    189 ACAAACTGGGGAGACACCTTACATGTACACTGAGTGGGGAGACAGCTTTGGCAGTATGTC 248
```

```
Qy     974 agtcctcatcaaaaacccaaggacacactctgggggaaagccttatgtgtgcagggaatg 1033
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     249 AGTCCTCATCAAAAACCCAAGGACACACTCTGGGGGAAAGCCTTATGTGTGCAGGGAATG 308

Qy    1034 tgggcgaggctttacgtggaagtcaaacctgatcacacatcagaggacacactcagggga 1093
           |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db     309 TGGGCGAGGCTTTACGTGGAAGTCAAACCTGATCACACACCAGAGGACACACTCAGGGGA 368

Qy    1094 gaaaccttatgtgtgcaaggattgtggacgaggctttacttggaagtcgaacctctttac 1153
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db     369 GAAACCTTATGTGTGCAAGGATTGTGGACGAGGCTTTACTTGGAAGTCGAACCTCCTTAC 428

Qy    1154 acatcagcggacacactcagggctcaagccttatgtgtgcaaggaatgtgggcagagctt 1213
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     429 ACATCAGCGGACACACTCAGGGCTCAAGCCTTATGTGTGCAAGGAATGTGGGCAGAGCTT 488

Qy    1214 tagcctgaagtcaaacctcattacccaccagagggcgcacactggggagaagccttatgt 1273
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     489 TAGCCTGAAGTCAAACCTCATTACCCACCAGAGGGCGCACACTGGGGAGAAGCCTTATGT 548

Qy    1274 ttgcagggaatgtgggcgtggctttcgccagca 1306
           |||||||||||||||||||||||||||||||
Db     549 TTGCAGGGAATGTGGGCGTGGCTTTCGCCAGCA 581


RESULT 11
BF111754/c
LOCUS       BF111754                512 bp    mRNA    linear   EST 20-OCT-2000
DEFINITION  7135a10.x1 Soares_NSF_F8_9W_OT_PA_P_S1 Homo sapiens cDNA clone
            IMAGE:3523123 3', mRNA sequence.
ACCESSION   BF111754
VERSION     BF111754.1  GI:10941444
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 512)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40UP from Gibco
            High quality sequence stop: 447.
FEATURES             Location/Qualifiers
     source          1. .512
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3523123"
                     /clone_lib="Soares_NSF_F8_9W_OT_PA_P_S1"
                     /lab_host="DH10B"
                     /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Equal amounts of plasmid DNA from five normalized
                     libraries were mixed, and ss circles were made in vitro.
                     Following HAP purification, this DNA was used as tracer in
                     a subtractive hybridization reaction. The driver was
                     PCR-amplified cDNAs from pools of 5,000 clones made from
                     the same 5 libraries. The pools consisted of the following
                     libraries and cloneIDs: Soares NbHSF pool 1:
                     309384-310919, 323208-325895 Soares Nb2HP pool 1:
                     145032-147335, 147720-148103, 148872-149255, 15002 -
                     150407, 151176-152327 Soares Nb2HF8-9W pool 1:
                     758280-760583, 772104-774407 Soares NbHPA pool 1:
                     304776-306311, 320136-322823, 326280-326663 Soares NbHOT
                     pool 1:  723720-726407, 739080-740999 Subtraction by Bento
                     Soares and M. Fatima Bonaldo."
BASE COUNT      144 a    111 c    140 g    117 t
```

```
ORIGIN


  Query Match           16.5%;  Score 458;  DB 10;  Length 512;
  Best Local Similarity   99.8%;  Pred. No. 6.6e-230;
  Matches  508;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy    2260 tgtatctccatccacctgaaggagaattgctggctcattttcaggagccctgcccttcct 2319
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     512 TGTATCTCCATCCACCTGAAGGAGAATTGCTGGCTCATTTTCAGGAGCCCTGCCCTTCCT 453

Qy    2320 cactgtggatggtgggttgtggaaacccggtcaggtaatgatagtggcaggaggcagtca 2379
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     452 CACTGTGGATGGTGGGTTGTGGAAACCCGGTCAGGTAATGATAGTGGCAGGAGGCAGTCA 393

Qy    2380 aatgcccaggcagataggggtgggtacctggtgaaacccaaccttaaagctgaagacagt 2439
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     392 AATGCCCAGGCAGATAGGGGTGGGTACCTGGTGAAACCCAACCTTAAAGCTGAAGACAGT 333

Qy    2440 cccggctaaatcctcatactgaattgagaacctgtcttcccatttggtgtgctttcctcc 2499
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     332 CCCGGCTAAATCCTCATACTGAATTGAGAACCTGTCTTCCCATTTGGTGTGCTTTCCTCC 273

Qy    2500 gattgatcccaaccccttcacctattttacgtatacctgccctttcctaattggtttttac 2559
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     272 GATTGATCCCAACCCTTCACCTATTTTACGTATACCTGCCCTTTCCTAATTGGTTTTTAC 213

Qy    2560 actgctgtgcccacctttgagtggtgcctttgcatacttacaaatcagtcaacgtgtat 2619
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     212 ACTGCTGTGCCCACCTTTGAGTGGTGCCTTTGCATACTTACAAATCAGTCAACGTGTAT 153

Qy    2620 tcccctattctgagcccataaaagacccagactcagctgcagtgaggagagaaatcaccc 2679
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     152 TCCCCTATTCTGAGCCCATAAAAGACCCAGACTCAGCTGCAGTGAGGAGAGAAATCACCC 93

Qy    2680 tgctgtggaggttggggaccactccctgcatcccctctccactgagagctgttcttttgc 2739
           |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      92 TGCTGTGGGGGTTGGGGACCACTCCCTGCATCCCCTCTCCACTGAGAGCTGTTCTTTTGC 33

Qy    2740 tcaataaaattcttttctacccatcctca 2768
           ||||||||||||||||||||||||||||||
Db      32 TCAATAAAATTCTTTTCTACCCATCCTCA 4


RESULT 12
BE463501/c
LOCUS       BE463501                505 bp    mRNA    linear   EST 27-JUL-2000
DEFINITION  hw24d02.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:3183843 3',
            mRNA sequence.
ACCESSION   BE463501
VERSION     BE463501.1  GI:9509274
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 505)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL, send email to:
            info@image.llnl.gov
            Seq primer: -40UP from Gibco
            High quality sequence stop: 455.
```

```
FEATURES             Location/Qualifiers
     source          1. .505
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3183843"
                     /clone_lib="NCI_CGAP_Kid11"
                     /lab_host="DH10B"
                     /note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Plasmid DNA from the normalized library NCI_CGAP_Kid3 was
                     prepared, and ss circles were made in vitro. Following HAP
                     purification, this DNA was used as tracer in a subtractive
                     hybridization reaction. The driver was PCR-amplified cDNAs
                     from a pool of 5,000 clones made from the same library
                     (cloneIDs 1322376-1323911, 1456007-1456775, and
                     1500552-1502855). Subtraction by Bento Soares and M.
                     Fatima Bonaldo.  "
BASE COUNT      142 a    110 c    140 g    113 t
ORIGIN


  Query Match             16.4%;  Score 454;  DB 10;  Length 505;
  Best Local Similarity   99.8%;  Pred. No. 8.4e-228;
  Matches  504;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy   2263 atctccatccacctgaaggagaattgctggctcattttcaggagccctgcccttcctcac 2322
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    505 ATCTCCATCCACCTGAAGGAGAATTGCTGGCTCATTTTCAGGAGCCCTGCCCTTCCTCAC 446

Qy   2323 tgtggatggtgggttgtggaaacccggtcaggtaatgatagtggcaggaggcagtcaaat 2382
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    445 TGTGGATGGTGGGTTGTGGAAACCCGGTCAGGTAATGATAGTGGCAGGAGGCAGTCAAAT 386

Qy   2383 gcccaggcagataggggtgggtacctggtgaaacccaaccttaaagctgaagacagtccc 2442
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    385 GCCCAGGCAGATAGGGGTGGGTACCTGGTGAAACCCAACCTTAAAGCTGAAGACAGTCCC 326

Qy   2443 ggctaaatcctcatactgaattgagaacctgtcttcccatttggtgtgctttcctccgat 2502
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    325 GGCTAAATCCTCATACTGAATTGAGAACCTGTCTTCCCATTTGGTGTGCTTTCCTCCGAT 266

Qy   2503 tgatcccaacccttcacctattttacgtatacctgccctttcctaattggtttttacact 2562
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    265 TGATCCCAACCCTTCACCTATTTTACGTATACCTGCCCTTTCCTAATTGGTTTTTACACT 206

Qy   2563 gctgtgcccacctttgagtggtgcctttgcatacttacaaatcagtcaacgtgtattcc 2622
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    205 GCTGTGCCCACCTTTTGAGTGGTGCCTTTGCATACTTACAAATCAGTCAACGTGTATTCC 146

Qy   2623 cctattctgagcccataaaagacccagactcagctgcagtgaggagagaaatcaccctgc 2682
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    145 CCTATTCTGAGCCCATAAAAGACCCAGACTCAGCTGCAGTGAGGAGAGAAATCACCCTGC 86

Qy   2683 tgtggaggttggggaccactccctgcatcccctctccactgagagctgttctttttgctca 2742
           ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     85 TGTGGGGGTTGGGGACCACTCCCTGCATCCCCTCTCCACTGAGAGCTGTTCTTTTGCTCA 26

Qy   2743 ataaaattcttttctacccatcctc 2767
           |||||||||||||||||||||||||
Db     25 ATAAAATTCTTTTCTACCCATCCTC 1


RESULT  13
AW401692
LOCUS       AW401692                486 bp    mRNA    linear   EST 16-FEB-2000
DEFINITION  UI-HF-BK0-aaf-c-03-0-UI.r1 NIH_MGC_36 Homo sapiens cDNA clone
            IMAGE:3053549 5', mRNA sequence.
ACCESSION   AW401692
VERSION     AW401692.1  GI:6920378
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 486)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Eco RI site shown at the beginning of the sequence.
             Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
             cDNA Library Preparation: M.B. Soares Lab
             cDNA Library Arrayed by: M.B. Soares Lab
             DNA Sequencing by: M.B. Soares Lab
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: M13 Forward.
FEATURES             Location/Qualifiers
     source          1. .486
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3053549"
                     /clone_lib="NIH_MGC_36"
                     /tissue_type="lymph"
                     /cell_type="germinal center B cells"
                     /cell_line="MGC85"
                     /lab_host="DH10B (LTI)"
                     /note="Vector: pT7T3-Pac; Site_1: NotI; Site_2: Eco RI;
                     Constructed from size fractionated cytoplasmic mRNA
                     (0.5-1.5kb). Directionally cloned. Cells provided by Louis
                     M. Staudt, Ph.D. Library preparation by Maria de Fatima
                     Bonaldo, Ph.D.  and M. Bento Soares, Ph.D."
BASE COUNT      138 a    113 c    136 g     98 t      1 others
ORIGIN


  Query Match             16.2%;  Score 449;  DB 9;  Length 486;
  Best Local Similarity   100.0%;  Pred. No. 3.6e-225;
  Matches  449;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    963 ggcagtatgtcagtcctcatcaaaaacccaaggacacactctgggggaaagccttatgtg 1022
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      8 GGCAGTATGTCAGTCCTCATCAAAAACCCAAGGACACACTCTGGGGGAAAGCCTTATGTG 67

Qy   1023 tgcagggaatgtgggcgaggctttacgtggaagtcaaacctgatcacacatcagaggaca 1082
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     68 TGCAGGGAATGTGGGCGAGGCTTTACGTGGAAGTCAAACCTGATCACACATCAGAGGACA 127

Qy   1083 cactcaggggagaaaccttatgtgtgcaaggattgtggacgaggctttacttggaagtcg 1142
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    128 CACTCAGGGGAGAAACCTTATGTGTGCAAGGATTGTGGACGAGGCTTTACTTGGAAGTCG 187

Qy   1143 aacctctttacacatcagcggacacactcagggctcaagccttatgtgtgcaaggaatgt 1202
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    188 AACCTCTTTACACATCAGCGGACACACTCAGGGCTCAAGCCTTATGTGTGCAAGGAATGT 247

Qy   1203 gggcagagctttagcctgaagtcaaacctcattacccaccagagggcgcacactggggag 1262
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    248 GGGCAGAGCTTTAGCCTGAAGTCAAACCTCATTACCCACCAGAGGGCGCACACTGGGGAG 307

Qy   1263 aagccttatgtttgcagggaatgtgggcgtggctttcgccagcattcacacctggtcaga 1322
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    308 AAGCCTTATGTTTGCAGGGAATGTGGGCGTGGCTTTCGCCAGCATTCACACCTGGTCAGA 367

Qy   1323 cacaagaggacacattcaggagagaagccttacatttgcagggagtgtgagcaaggcttt 1382
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    368 CACAAGAGGACACATTCAGGAGAGAAGCCTTACATTTGCAGGGAGTGTGAGCAAGGCTTT 427

Qy   1383 agccagaagtcacacctcatcagacactt 1411
           |||||||||||||||||||||||||||||
Db    428 AGCCAGAAGTCACACCTCATCAGACACTT 456


RESULT  14
```

```
AV720986
LOCUS       AV720986                486 bp    mRNA    linear   EST 16-OCT-2000
DEFINITION  AV720986 HTB Homo sapiens cDNA clone HTBBIC05 5', mRNA sequence.
ACCESSION   AV720986
VERSION     AV720986.1  GI:10818138
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 486)
  AUTHORS   Gu,Y., Peng,Y., Song,H., Huang,Q., Yang,Y., Gao,G., Xiao,H., Xu,X.,
            Li,N., Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng,L., Xu
            ,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Hu,R.,
            Chen,J., Chen,Z. and Han,Z.
  TITLE     Homo sapiens cDNA HTB clones
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Zeguang Han
            Chinese National Human Genome Center at Shanghai
            351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
            201203, P. R. China
            Tel: 86-21-50801919(ex.45)
            Fax: 86-21-50801922
            Email: hanzg@chgc.sh.cn
            This clone is available at CHGC in Shanghai.
FEATURES             Location/Qualifiers
     source          1. .486
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="HTBBIC05"
                     /clone_lib="HTB"
                     /tissue_type="Hypothalamus"
                     /dev_stage="Adult"
                     /lab_host="SOLR"
                     /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
                     XhoI"
BASE COUNT      104 a    134 c    110 g    137 t      1 others
ORIGIN
```

```
  Query Match             15.1%;  Score 419;  DB 9;  Length 486;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-209;
  Matches  419;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    2269 atccacctgaaggagaattgctggctcattttcaggagccctgcccttcctcactgtgga 2328
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATCCACCTGAAGGAGAATTGCTGGCTCATTTTCAGGAGCCCTGCCCTTCCTCACTGTGGA 60

Qy    2329 tggtgggttgtggaaacccggtcaggtaatgatagtggcaggaggcagtcaaatgcccag 2388
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TGGTGGGTTGTGGAAACCCGGTCAGGTAATGATAGTGGCAGGAGGCAGTCAAATGCCCAG 120

Qy    2389 gcagataggggtgggtacctggtgaaacccaaccttaaagctgaagacagtcccggctaa 2448
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GCAGATAGGGGTGGGTACCTGGTGAAACCCAACCTTAAAGCTGAAGACAGTCCCGGCTAA 180

Qy    2449 atcctcatactgaattgagaacctgtcttcccatttggtgtgctttcctccgattgatcc 2508
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 ATCCTCATACTGAATTGAGAACCTGTCTTCCCATTTGGTGTGCTTTCCTCCGATTGATCC 240

Qy    2509 caacccttcacctattttacgtatacctgccctttcctaattggtttttacactgctgtg 2568
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CAACCCTTCACCTATTTTACGTATACCTGCCCTTTCCTAATTGGTTTTTACACTGCTGTG 300

Qy    2569 cccacctttgagtggtgcctttgcatacttacaaatcagtcaacgtgtattcccctatt 2628
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CCCACCTTTTGAGTGGTGCCTTTGCATACTTACAAATCAGTCAACGTGTATTCCCCTATT 360

Qy    2629 ctgagcccataaaagacccagactcagctgcagtgaggagagaaatcaccctgctgtgg 2687
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 CTGAGCCCATAAAAGACCCAGACTCAGCTGCAGTGAGGAGAGAAATCACCCTGCTGTGG 419
```

```
RESULT  15
AW964787/c
LOCUS       AW964787                613 bp    mRNA    linear   EST 01-JUN-2000
DEFINITION  EST376980 MAGE resequences, MAGH Homo sapiens cDNA, mRNA sequence.
ACCESSION   AW964787
VERSION     AW964787.1  GI:8154743
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 613)
  AUTHORS   Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspard,R., Gay,C., Holt
            ,I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and
            Quackenbush,J.
  TITLE     Assessment of gene expression patterns in a model of colon tumor
            metastasis using a 19,200 element cDNA microarray
  JOURNAL   Unpublished (2000)
COMMENT     Contact: John Quackenbush
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 3528
            Fax: 301 838 0208
            Email: johnq@tigr.org
            Plate: 207
            Seq primer: Forward.
FEATURES             Location/Qualifiers
     source          1. .613
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="MAGE resequences, MAGH"
                     /note="Vector: pBluescriptSKm"
BASE COUNT      171 a    140 c    158 g    144 t
ORIGIN
```

```
  Query Match             14.9%;  Score 414;  DB 9;  Length 613;
  Best Local Similarity   99.8%;  Pred. No. 1.1e-206;
  Matches  464;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy    2223 tgtgtgtgattatgcatgagactgtactggtaagacttgtatctccatccacctgaagga 2282
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     544 TGTGTGTGATTATGCATGAGACTGTACTGGTAAGACTTGTATCTCCATCCACCTGAAGGA 485

Qy    2283 gaattgctggctcattttcaggagccctgcccttcctcactgtggatggtgggttgtgga 2342
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     484 GAATTGCTGGCTCATTTTCAGGAGCCCTGCCCTTCCTCACTGTGGATGGTGGGTTGTGGA 425

Qy    2343 aacccggtcaggtaatgatagtggcaggaggcagtcaaatgcccaggcagataggggtgg 2402
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     424 AACCCGGTCAGGTAATGATAGTGGCAGGAGGCAGTCAAATGCCCAGGCAGATAGGGGTGG 365

Qy    2403 gtacctggtgaaacccaaccttaaagctgaagacagtcccggctaaatcctcatactgaa 2462
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     364 GTACCTGGTGAAACCCAACCTTAAAGCTGAAGACAGTCCCGGCTAAATCCTCATACTGAA 305

Qy    2463 ttgagaacctgtcttcccatttggtgtgctttcctccgattgatcccaacccttcaccta 2522
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     304 TTGAGAACCTGTCTTCCCATTTGGTGTGCTTTCCTCCGATTGATCCCAACCCTTCACCTA 245

Qy    2523 ttttacgtatacctgccctttcctaattggtttttacactgctgtgcccaccttttgagt 2582
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     244 TTTTACGTATACCTGCCCTTTCCTAATTGGTTTTTACACTGCTGTGCCCACCTTTTGAGT 185

Qy    2583 ggtgcctttgcatacttacaaatcagtcaacgtgtattcccctattctgagcccataaaa 2642
           |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db     184 GGTGCCTTTGCATACTTACAAATCAGTCAACGTGTATTCCCCTATTTTGAGCCCATAAAA 125

Qy    2643 gacccagactcagctgcagtgaggagagaaatcaccctgctgtgg 2687
           |||||||||||||||||||||||||||||||||||||||||||||
Db     124 GACCCAGACTCAGCTGCAGTGAGGAGAGAAATCACCCTGCTGTGG 80
```

Search completed: May 17, 2002, 09:22:49
Job time: 71046 sec

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: May 17, 2002, 13:55:18 ; Search time 105.07 Seconds
(without alignments)
6480.404 Million cell updates/sec

Title: US-09-898-556A-3
Perfect score: 2772
Sequence: 1 caggcgcgttaagctggttg.....tttctacccatcctcaccct 2772

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 383533 seqs, 122816752 residues

Word size : 0

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	21	0.8	2382	4	US-09-588-256-9	Sequence 9, Appli
	2	20	0.7	1892	2	US-08-933-750C-66	Sequence 66, Appl
	3	20	0.7	1892	3	US-09-234-613-66	Sequence 66, Appl
	4	19	0.7	3240	4	US-09-262-773-7	Sequence 7, Appli
	5	19	0.7	3244	4	US-09-262-773-3	Sequence 3, Appli
	6	19	0.7	3264	4	US-09-262-773-5	Sequence 5, Appli
	7	19	0.7	3268	4	US-09-262-773-1	Sequence 1, Appli
	8	19	0.7	3810	2	US-08-475-844-8	Sequence 8, Appli
	9	19	0.7	3810	5	PCT-US95-08429-8	Sequence 8, Appli
	10	19	0.7	20137	4	US-09-262-773-206	Sequence 206, App
	11	19	0.7	20138	4	US-09-262-773-9	Sequence 9, Appli
	12	19	0.7	23071	4	US-09-262-773-210	Sequence 210, App
c	13	19	0.7	246240	2	US-08-724-394A-20	Sequence 20, Appl
c	14	19	0.7	246240	2	US-08-724-394A-21	Sequence 21, Appl
c	15	19	0.7	246240	2	US-08-724-394A-22	Sequence 22, Appl
	16	18	0.6	575	1	US-08-507-016-8	Sequence 8, Appli
c	17	18	0.6	1027	4	US-09-303-524A-1	Sequence 1, Appli
	18	18	0.6	1558	1	US-08-416-870C-9	Sequence 9, Appli
	19	18	0.6	2133	2	US-08-820-170A-11	Sequence 11, Appl
	20	18	0.6	2133	3	US-09-055-699-11	Sequence 11, Appl
	21	18	0.6	2133	4	US-09-273-565-11	Sequence 11, Appl
	22	18	0.6	2133	4	US-09-565-538-11	Sequence 11, Appl
	23	18	0.6	3754	2	US-08-820-170A-12	Sequence 12, Appl
	24	18	0.6	3754	3	US-09-055-699-12	Sequence 12, Appl
	25	18	0.6	3754	4	US-09-273-565-12	Sequence 12, Appl
	26	18	0.6	3754	4	US-09-565-538-12	Sequence 12, Appl
c	27	18	0.6	7886	2	US-08-751-189-2	Sequence 2, Appli
c	28	18	0.6	7886	2	US-09-060-836-2	Sequence 2, Appli
c	29	18	0.6	7886	4	US-09-184-445-2	Sequence 2, Appli
	30	17	0.6	80	5	PCT-US91-02942-99	Sequence 99, Appl
c	31	17	0.6	273	3	US-08-956-182-21	Sequence 21, Appl
c	32	17	0.6	289	1	US-08-592-126-124	Sequence 124, App
	33	17	0.6	347	4	US-08-905-223-133	Sequence 133, App
	34	17	0.6	532	4	US-09-328-111-648	Sequence 648, App
c	35	17	0.6	563	4	US-09-385-982-433	Sequence 433, App
c	36	17	0.6	732	3	US-08-956-182-16	Sequence 16, Appl
	37	17	0.6	1362	4	US-09-171-461-47	Sequence 47, Appl
c	38	17	0.6	1561	1	US-07-662-007B-38	Sequence 38, Appl
c	39	17	0.6	1561	1	US-07-968-971A-1	Sequence 1, Appli
c	40	17	0.6	1561	1	US-07-824-247-38	Sequence 38, Appl
c	41	17	0.6	1561	1	US-07-824-247-41	Sequence 41, Appl
c	42	17	0.6	1561	1	US-08-142-473A-1	Sequence 1, Appli
c	43	17	0.6	1561	1	US-08-424-406-1	Sequence 1, Appli
c	44	17	0.6	1561	1	US-08-464-523B-5	Sequence 5, Appli
c	45	17	0.6	1561	1	US-08-469-203A-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-588-256-9
; Sequence 9, Application US/09588256
; Patent No. 6291665
; GENERAL INFORMATION:
; APPLICANT: Gaffney, Thomas
; APPLICANT: Flavier, Albert
; APPLICANT: Gates, Krista
; APPLICANT: Wendland, Juergen
; APPLICANT: Ayad-Durieux, Yasmina
; APPLICANT: Dietrich, Fred
; APPLICANT: Philippsen, Peter
; TITLE OF INVENTION: Fungal Target Genes and Methods
; FILE REFERENCE: PB/5-30908A
; CURRENT APPLICATION NUMBER: US/09/588,256
; CURRENT FILING DATE: 2000-06-06
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 2382
; TYPE: DNA
; ORGANISM: Ashbya gossypii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2382)
US-09-588-256-9

Query Match 0.8%; Score 21; DB 4; Length 2382;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1875 ggcttttgtgctaagttaact 1895
||||||||||||||||||||
Db 1990 ggcttttgtgctaagttaact 2010

RESULT 2
US-08-933-750C-66
; Sequence 66, Application US/08933750C
; Patent No. 5932442
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.

```
;    TITLE OF INVENTION:  HUMAN REGULATORY MOLECULES
;    NUMBER OF SEQUENCES:  98
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/933,750C
;      FILING DATE:  September 23, 1997
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0356 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  66:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1892 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  BRSTNOT03
;      CLONE:  641127
US-08-933-750C-66


  Query Match             0.7%;  Score 20;  DB 2;  Length 1892;
  Best Local Similarity   100.0%;  Pred. No. 3.4;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1419 cacacaggagagaagcctta 1438
          ||||||||||||||||||||
Db   1289 CACACAGGAGAGAAGCCTTA 1308


RESULT   3
US-09-234-613-66
; Sequence 66, Application US/09234613
; Patent No. 6132973
;  GENERAL INFORMATION:
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Shah, Purvi
;    APPLICANT:  Au-Young, Janice
;    APPLICANT:  Yue, Henry
;    APPLICANT:  Guegler, Karl J.
;    APPLICANT:  Corley, Neil C.
;    TITLE OF INVENTION:  HUMAN REGULATORY MOLECULES
;    NUMBER OF SEQUENCES:  98
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/234,613
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/933,750
;      FILING DATE:  September 23, 1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0356 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  66:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1892 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  BRSTNOT03
;      CLONE:  641127
US-09-234-613-66


  Query Match             0.7%;  Score 20;  DB 3;  Length 1892;
  Best Local Similarity   100.0%;  Pred. No. 3.4;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1419 cacacaggagagaagcctta 1438
          ||||||||||||||||||||
Db   1289 CACACAGGAGAGAAGCCTTA 1308


RESULT   4
US-09-262-773-7
; Sequence 7, Application US/09262773
; Patent No. 6225451
; GENERAL INFORMATION:
;  APPLICANT: Ballinger, Dennis G.
;  APPLICANT: Ding, Wei
;  APPLICANT: Wagner, Susanne
;  APPLICANT: Hess, Mark A.
;  TITLE OF INVENTION: CHROMOSOME 11-LINKED CORONARY HEART DISEASE
;  TITLE OF INVENTION: SUSCEPTIBILITY GENE CHD1
;  FILE REFERENCE: Myriad 3
;  CURRENT APPLICATION NUMBER: US/09/262,773
;  CURRENT FILING DATE: 1999-03-04
;  NUMBER OF SEQ ID NOS: 210
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
;   LENGTH: 3240
;   TYPE: DNA
;   ORGANISM: human
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (98)..(523)
US-09-262-773-7


  Query Match             0.7%;  Score 19;  DB 4;  Length 3240;
  Best Local Similarity   100.0%;  Pred. No. 11;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   2068 tcagacatcagaggacaca 2086
          ||||||||||||||||||||
Db   1559 tcagacatcagaggacaca 1577
```

```
RESULT   5
US-09-262-773-3
; Sequence 3, Application US/09262773
; Patent No. 6225451
; GENERAL INFORMATION:
;  APPLICANT: Ballinger, Dennis G.
;  APPLICANT: Ding, Wei
;  APPLICANT: Wagner, Susanne
;  APPLICANT: Hess, Mark A.
;  TITLE OF INVENTION: CHROMOSOME 11-LINKED CORONARY HEART DISEASE
;  TITLE OF INVENTION: SUSCEPTIBILITY GENE CHD1
;  FILE REFERENCE: Myriad 3
;  CURRENT APPLICATION NUMBER: US/09/262,773
;  CURRENT FILING DATE: 1999-03-04
;  NUMBER OF SEQ ID NOS: 210
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 3244
;   TYPE: DNA
;   ORGANISM: human
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (98)..(2017)
US-09-262-773-3


  Query Match             0.7%;  Score 19;  DB 4;  Length 3244;
  Best Local Similarity   100.0%;  Pred. No. 11;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   2068 tcagacatcagaggacaca 2086
           ||||||||||||||||||||
Db   1563 tcagacatcagaggacaca 1581


RESULT   6
US-09-262-773-5
; Sequence 5, Application US/09262773
; Patent No. 6225451
; GENERAL INFORMATION:
;  APPLICANT: Ballinger, Dennis G.
;  APPLICANT: Ding, Wei
;  APPLICANT: Wagner, Susanne
;  APPLICANT: Hess, Mark A.
;  TITLE OF INVENTION: CHROMOSOME 11-LINKED CORONARY HEART DISEASE
;  TITLE OF INVENTION: SUSCEPTIBILITY GENE CHD1
;  FILE REFERENCE: Myriad 3
;  CURRENT APPLICATION NUMBER: US/09/262,773
;  CURRENT FILING DATE: 1999-03-04
;  NUMBER OF SEQ ID NOS: 210
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
;   LENGTH: 3264
;   TYPE: DNA
;   ORGANISM: human
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (98)..(496)
US-09-262-773-5


  Query Match             0.7%;  Score 19;  DB 4;  Length 3264;
  Best Local Similarity   100.0%;  Pred. No. 11;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   2068 tcagacatcagaggacaca 2086
           ||||||||||||||||||||
Db   1583 tcagacatcagaggacaca 1601


RESULT   7
US-09-262-773-1
; Sequence 1, Application US/09262773
; Patent No. 6225451
; GENERAL INFORMATION:
;  APPLICANT: Ballinger, Dennis G.
;  APPLICANT: Ding, Wei
;  APPLICANT: Wagner, Susanne
;  APPLICANT: Hess, Mark A.
;  TITLE OF INVENTION: CHROMOSOME 11-LINKED CORONARY HEART DISEASE
;  TITLE OF INVENTION: SUSCEPTIBILITY GENE CHD1
;  FILE REFERENCE: Myriad 3
;  CURRENT APPLICATION NUMBER: US/09/262,773
;  CURRENT FILING DATE: 1999-03-04
;  NUMBER OF SEQ ID NOS: 210
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 3268
;   TYPE: DNA
;   ORGANISM: human
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (98)..(2041)
US-09-262-773-1


  Query Match             0.7%;  Score 19;  DB 4;  Length 3268;
  Best Local Similarity   100.0%;  Pred. No. 11;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   2068 tcagacatcagaggacaca 2086
           ||||||||||||||||||||
Db   1587 tcagacatcagaggacaca 1605


RESULT   8
US-08-475-844-8
; Sequence 8, Application US/08475844
; Patent No. 5972643
;  GENERAL INFORMATION:
;    APPLICANT:  Lobanenkov, Victor V.
;    APPLICANT:  Neiman, Paul E.
;    APPLICANT:  Klenova, Elena M.
;    APPLICANT:  Goodwin, Graham H.
;    APPLICANT:  Filippova, Galina N.
;    APPLICANT:  Collins, Steven J.
;    TITLE OF INVENTION:  CTCF
;    NUMBER OF SEQUENCES:  21
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend Khourie and Crew
;      STREET:  One Market Plaza
;      CITY:  San Francisco
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/475,844
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/261,680
;      FILING DATE:  17-JUN-1994
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Parmelee, Steven W.
;      REGISTRATION NUMBER:  31,990
;      REFERENCE/DOCKET NUMBER:  14538A-11-1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  206-467-9600
```

```
;      TELEFAX:  415-543-5043
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3810 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
;    IMMEDIATE SOURCE:
;      CLONE:  human CTCF cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  292..2475
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  281..1074
;      OTHER INFORMATION:  /label= exon2
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  1075..1245
;      OTHER INFORMATION:  /label= exon3
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  1246..1379
;      OTHER INFORMATION:  /label= exon4
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  1380..1499
;      OTHER INFORMATION:  /label= exon5
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  1500..1649
;      OTHER INFORMATION:  /label= exon6
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  1650..1810
;      OTHER INFORMATION:  /label= exon7
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  1810..1992
;      OTHER INFORMATION:  /label= exon8
US-08-475-844-8


  Query Match             0.7%;  Score 19;  DB 2;  Length 3810;
  Best Local Similarity   100.0%;  Pred. No. 10;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1755 cattcaggggaaaagcctt 1773
          ||||||||||||||||||
Db   1492 CATTCAGGGGAAAAGCCTT 1510


RESULT   9
PCT-US95-08429-8
; Sequence 8, Application PC/TUS9508429
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  CTCF
;    NUMBER OF SEQUENCES:  21
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/08429
;      FILING DATE:  15-JUN-1995
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/261,680
;      FILING DATE:  17-JUN-1994
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Parmelee, Steven W.
;      REGISTRATION NUMBER:  31,990
;      REFERENCE/DOCKET NUMBER:  14538A-11-1PC
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  206-467-9600
;      TELEFAX:  415-543-5043
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3810 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
;    IMMEDIATE SOURCE:
;      CLONE:  human CTCF cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  292..2475
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  281..1074
;      OTHER INFORMATION:  /label= exon2
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  1075..1245
;      OTHER INFORMATION:  /label= exon3
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  1246..1379
;      OTHER INFORMATION:  /label= exon4
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  1380..1499
;      OTHER INFORMATION:  /label= exon5
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  1500..1649
;      OTHER INFORMATION:  /label= exon6
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  1650..1810
;      OTHER INFORMATION:  /label= exon7
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  1810..1992
;      OTHER INFORMATION:  /label= exon8
PCT-US95-08429-8


  Query Match             0.7%;  Score 19;  DB 5;  Length 3810;
  Best Local Similarity   100.0%;  Pred. No. 10;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1755 cattcaggggaaaagcctt 1773
          ||||||||||||||||||
Db   1492 CATTCAGGGGAAAAGCCTT 1510


RESULT  10
US-09-262-773-206
; Sequence 206, Application US/09262773
; Patent No. 6225451
; GENERAL INFORMATION:
;  APPLICANT: Ballinger, Dennis G.
;  APPLICANT: Ding, Wei
;  APPLICANT: Wagner, Susanne
;  APPLICANT: Hess, Mark A.
;  TITLE OF INVENTION: CHROMOSOME 11-LINKED CORONARY HEART DISEASE
```

```
;  TITLE OF INVENTION: SUSCEPTIBILITY GENE CHD1
;  FILE REFERENCE: Myriad 3
;  CURRENT APPLICATION NUMBER: US/09/262,773
;  CURRENT FILING DATE: 1999-03-04
;  NUMBER OF SEQ ID NOS: 210
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 206
;    LENGTH: 20137
;    TYPE: DNA
;    ORGANISM: human
US-09-262-773-206


  Query Match             0.7%;  Score 19;  DB 4;  Length 20137;
  Best Local Similarity   100.0%;  Pred. No. 10;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    2068 tcagacatcagaggacaca 2086
           ||||||||||||||||||||
Db   14680 tcagacatcagaggacaca 14698


RESULT  11
US-09-262-773-9
; Sequence 9, Application US/09262773
; Patent No. 6225451
; GENERAL INFORMATION:
;  APPLICANT: Ballinger, Dennis G.
;  APPLICANT: Ding, Wei
;  APPLICANT: Wagner, Susanne
;  APPLICANT: Hess, Mark A.
;  TITLE OF INVENTION: CHROMOSOME 11-LINKED CORONARY HEART DISEASE
;  TITLE OF INVENTION: SUSCEPTIBILITY GENE CHD1
;  FILE REFERENCE: Myriad 3
;  CURRENT APPLICATION NUMBER: US/09/262,773
;  CURRENT FILING DATE: 1999-03-04
;  NUMBER OF SEQ ID NOS: 210
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
;    LENGTH: 20138
;    TYPE: DNA
;    ORGANISM: human
US-09-262-773-9


  Query Match             0.7%;  Score 19;  DB 4;  Length 20138;
  Best Local Similarity   100.0%;  Pred. No. 10;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    2068 tcagacatcagaggacaca 2086
           ||||||||||||||||||||
Db   14681 tcagacatcagaggacaca 14699


RESULT  12
US-09-262-773-210
; Sequence 210, Application US/09262773
; Patent No. 6225451
; GENERAL INFORMATION:
;  APPLICANT: Ballinger, Dennis G.
;  APPLICANT: Ding, Wei
;  APPLICANT: Wagner, Susanne
;  APPLICANT: Hess, Mark A.
;  TITLE OF INVENTION: CHROMOSOME 11-LINKED CORONARY HEART DISEASE
;  TITLE OF INVENTION: SUSCEPTIBILITY GENE CHD1
;  FILE REFERENCE: Myriad 3
;  CURRENT APPLICATION NUMBER: US/09/262,773
;  CURRENT FILING DATE: 1999-03-04
;  NUMBER OF SEQ ID NOS: 210
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 210
;    LENGTH: 23071
;    TYPE: DNA
;    ORGANISM: human
US-09-262-773-210


  Query Match             0.7%;  Score 19;  DB 4;  Length 23071;
  Best Local Similarity   100.0%;  Pred. No. 10;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    2068 tcagacatcagaggacaca 2086
           ||||||||||||||||||||
Db   17614 tcagacatcagaggacaca 17632


RESULT  13
US-08-724-394A-20/c
; Sequence 20, Application US/08724394A
; Patent No. 5872237
;  GENERAL INFORMATION:
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Kronmal, Gregory S.
;    APPLICANT:  Lauer, Peter M.
;    APPLICANT:  Ruddy, David A.
;    APPLICANT:  Thomas, Winston
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Megabase Transcript Map:  No. 5872237el
;    TITLE OF INVENTION:  Sequences and Antibodies Thereto
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  TOWNSEND and TOWNSEND and CREW LLP
;      STREET:  Two Embarcadero Center, 8th Floor
;      CITY:  San Francisco
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/724,394A
;      FILING DATE:  01-OCT-1996
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fitts, Renee A.
;      REGISTRATION NUMBER:  35,136
;      REFERENCE/DOCKET NUMBER:  017957-000100
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-576-0200
;      TELEFAX:  415-576-0300
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  246240 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..246240
;      OTHER INFORMATION:  /note= "HLA-H.CONTIG"
US-08-724-394A-20


  Query Match             0.7%;  Score 19;  DB 2;  Length 246240;
  Best Local Similarity   100.0%;  Pred. No. 10;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    2298 tttcaggagccctgcccttt 2316
           ||||||||||||||||||||
Db   36115 TTTCAGGAGCCCTGCCCTT 36097
```

```
RESULT  14
US-08-724-394A-21/c
; Sequence 21, Application US/08724394A
; Patent No. 5872237
;  GENERAL INFORMATION:
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Kronmal, Gregory S.
;    APPLICANT:  Lauer, Peter M.
;    APPLICANT:  Ruddy, David A.
;    APPLICANT:  Thomas, Winston
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Megabase Transcript Map:  No. 5872237el
;    TITLE OF INVENTION:  Sequences and Antibodies Thereto
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  TOWNSEND and TOWNSEND and CREW LLP
;      STREET:  Two Embarcadero Center, 8th Floor
;      CITY:  San Francisco
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/724,394A
;      FILING DATE:  01-OCT-1996
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fitts, Renee A.
;      REGISTRATION NUMBER:  35,136
;      REFERENCE/DOCKET NUMBER:  017957-000100
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-576-0200
;      TELEFAX:  415-576-0300
;  INFORMATION FOR SEQ ID NO:  21:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  246240 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..246240
;      OTHER INFORMATION:  /note= "HLA-H.CONTIG"
US-08-724-394A-21


  Query Match             0.7%;  Score 19;  DB 2;  Length 246240;
  Best Local Similarity   100.0%;  Pred. No. 10;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    2298 tttcaggagccctgccctt 2316
           ||||||||||||||||||
Db   36115 TTTCAGGAGCCCTGCCCTT 36097


RESULT  15
US-08-724-394A-22/c
; Sequence 22, Application US/08724394A
; Patent No. 5872237
;  GENERAL INFORMATION:
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Kronmal, Gregory S.
;    APPLICANT:  Lauer, Peter M.
;    APPLICANT:  Ruddy, David A.
;    APPLICANT:  Thomas, Winston
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Megabase Transcript Map:  No. 5872237el
;    TITLE OF INVENTION:  Sequences and Antibodies Thereto
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  TOWNSEND and TOWNSEND and CREW LLP
;      STREET:  Two Embarcadero Center, 8th Floor
;      CITY:  San Francisco
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/724,394A
;      FILING DATE:  01-OCT-1996
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fitts, Renee A.
;      REGISTRATION NUMBER:  35,136
;      REFERENCE/DOCKET NUMBER:  017957-000100
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-576-0200
;      TELEFAX:  415-576-0300
;  INFORMATION FOR SEQ ID NO:  22:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  246240 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..246240
;      OTHER INFORMATION:  /note= "HLA-H.CONTIG"
US-08-724-394A-22


  Query Match             0.7%;  Score 19;  DB 2;  Length 246240;
  Best Local Similarity   100.0%;  Pred. No. 10;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    2298 tttcaggagccctgccctt 2316
           ||||||||||||||||||
Db   36115 TTTCAGGAGCCCTGCCCTT 36097


Search completed: May 17, 2002, 16:27:42
Job time: 9144 sec
```